ANNALES

DE LA

SOCIÉTÉ SCIENTIFIQUE

DE BRUXELLES

ASSOCIATION SANS BUT LUCRATIF

TOME SOIXANTE-DIXIÈME, 1956

SÉRIE I

SCIENCES MATHÉMATIQUES, ASTRONOMIQUES ET PHYSIQUES

DEUXIÈME FASCICULE

17 AOUT 1956

Publié avec le concours de la Fondation universitaire de Belgique et du Gouvernement

LOUVAIN

Secrétariat de la Société Scientifique

11, RUE DES RÉCOLLETS, 11

Chèques postaux 2027.46

1956

Publication trimestrielle. Prix de ce fascicule séparé: 60 frs

AVIS

Les périodiques internationaux de références et de bibliographie sont autorisés à reproduire, moyennant due référence, les résumés que les auteurs placent en tête de leurs travaux dans ces *Annales*.

TABLE DE MATIÈRES

Session du 19 avril 1956, à Bruxelles — PAGES

PREMIÈRE SECTION : **Sciences mathématiques et astronomiques**

Thorème de Picard et structure des fonctions. Rectifications et compléments, par CH. DE LA VALLÉE POUSSIN 81

Produits scalaires à annulation symétrique, par R. BALLIEU 87

A propos de la notion de fonction aléatoire, par H. BRENY 96

Intégration par analyse harmonique, par G. LEMAITRE 117

Calcul des probabilités dans certains espaces topologiques, 2e partie, par R. LAVENDHOMME 124

DEUXIÈME SECTION : **Sciences physiques et chimiques** 164

Le prix d'ABONNEMENT aux ANNALES, série I, pour des personnes ne faisant pas partie de la Société scientifique, est fixé comme suit :

en Belgique, au Congo-Belge et au Luxembourg 100 fr.
dans les autres pays 150 fr.

SESSION DU 19 AVRIL 1956
A BRUXELLES

Première Section

Sciences Mathématiques et Astronomiques

Théorème de Picard et structure des fonctions Rectifications et compléments

PAR

Ch. de La VALLÉE POUSSIN

§ 1. — VALEURS EXCEPTIONNELLES

1. RECTIFICATIONS. — Dans un précédent article des *Annales* (T. LXIX, série I, 1955), nous avons proposé une généralisation de la définition de *valeur exceptionnelle*. Dans son sens initial, dû à Picard à l'occasion de son célèbre théorème, une *valeur exceptionnelle* est une valeur que la fonction $f(z)$ ne prend pas ou ne prend qu'un nombre limité de fois. Au sens généralisé nouveau, une valeur exceptionnelle est une valeur vers laquelle la fonction $f(z)$ converge quand z tend vers l'infini sur une ligne illimitée correspondante. La définition de Picard rentre comme cas particulier dans celle-ci.

Dans le même article, nous avons tenté une démonstration topologique du théorème de Picard qui, si elle était exacte, étendrait ce théorème aux valeurs exceptionnelles généralisées, dont le nombre ne pourrait surpasser 2. Nous avons reconnu que cette démonstration topologique est inexacte (n° 7 de l'article) et que l'extension du théorème est fausse. Nous en donnons la preuve par un exemple au n° suivant.

Dans ce même article, nous avons étudié la structure des fonctions par *tranches de période*, l'erreur précédente rend cette étude différente sauf si la fonction considérée admet et admet seulement deux valeurs exceptionnelles au sens strict de Picard. Nous reviendrons au § 2 sur ce mode de structure des fonctions.

2. Multiplicité des valeurs exceptionnelles. — Posons $z = x + yi$ et soit

$$u = f(z) = X + Yi$$

une fonction holomorphe de z admettant, au sens de Picard, les deux valeurs exceptionnelles 0 et ∞. La fonction holomorphe

$$u_1 = e^u$$

admet aussi les valeurs exceptionnelles 0 et ∞ au sens de Picard. Elle admet en plus, au sens généralisé, la valeur exceptionnelle 1 qui correspond à $u = 0$. Cette valeur est atteinte une infinité de fois. En effet, si z décrit la ligne $X = 0$, la fonction u_1 se réduit à

$$u_1 = e^{Yi}$$

qui est de module 1, et par conséquent, u décrit la circonférence de rayon 1 autour de l'origine. Quand u tourne indéfiniment sur cette circonférence, z décrit la ligne $X = 0$ et passe une infinité de fois par un nouveau représentant du point $u = 1$.

Ce résultat se généralise. On peut construire de proche en proche une fonction admettant, au sens généralisé, un nombre quelconque de valeurs exceptionnelles. En effet, soit u une fonction holomorphe admettant les valeurs exceptionnelles 0 et ∞ au sens de Picard, et n valeurs exceptionnelles $\alpha_1, \alpha_2, \ldots \alpha_n$ au sens généralisé. La fonction

$$u_1 = e^u$$

est holomorphe. Elle admet les valeurs 0, ∞ au sens de Picard, et au sens généralisé, les $n + 1$ valeurs exceptionnelles

$$1, e^{\alpha_1}, e^{\alpha_2}, \ldots e^{\alpha_n}$$

correspondant aux valeurs $0, \alpha_1, \alpha_2, \ldots \alpha_n$ de u. Ce procédé substitue à la valeur exceptionnelle 0 au sens de Picard la valeur exceptionnelle 1 au sens généralisé.

§ 2. — LA STRUCTURE PAR TRANCHES

3. Étude d'une fonction particulière. — Nous prenons comme exemple la fonction

$$u = e^{(e^z)}.$$

C'est le cas le plus simple des fonctions considérées au n° précé-

dent. Cette fonction holomorphe admet donc trois valeurs exceptionnelles, les valeurs 0 et ∞ au sens de Picard, et la valeur 1 au sens généralisé.

La fonction e^z est périodique de période $2\pi i$. Ses tranches de période sont limitées par les parallèles à l'axe des x d'équations $y = 2k\pi i$ où k est un entier positif ou négatif. La fonction u, qui reprend les mêmes valeurs en même temps que e^z admet donc les mêmes tranches de période. Nous appellerons celles-ci ses *tranches primaires* de période. Comme le comportement de la fonction est le même dans toutes ces tranches, il nous suffira de considérer l'une d'elles, celle comprise entre les valeurs 0 et 2π de y, que nous désignerons par T.

Nous allons montrer que la nouvelle valeur exceptionnelle a pour effet de déterminer une véritable prolifération des tranches de périodes par subdivisions.

Nous désignerons par A l'un ou l'autre des deux bords $y = 0$ ou $y = 2\pi$ de la tranche primaire T, et par B sa médiane $y = \pi$.

Quand z décrit A, x varie de $-\infty$ à $+\infty$, $e^z = e^x$ varie de 0 à ∞, et u de 1 à ∞.

Quand z décrit B, x varie de $-\infty$ à $+\infty$, $e^z = -e^x$ varie de 0 à $-\infty$ et u varie de 1 à 0.

A la ligne A du plan (z) correspond dans le plan (u) un segment a de l'axe réel allant du point 1 à l'infini.

A la ligne B du plan (z) correspond dans le plan (u) un segment b de l'axe réel joignant les points 0 et 1.

Chacune des deux moitiés de la tranche primaire T *séparées par la ligne* B, *est une tranche de période, c'est-à-dire que tous les points du plan* (u) *y sont représentés.*

En effet, les lignes A et B limitent une partie du plan (z), tandis que les lignes correspondantes a et b ne partagent pas le plan (u). Ce premier partage n'est qu'un point de départ.

Déterminons d'abord les lignes du plan (z) qui correspondent à la circonférence γ de rayon 1 autour de l'origine dans le plan (u). Posant $z = x + yi$, on a

$$e^z = e^x (\cos y + i \sin y).$$

Donc, pour $\cos y = 0$, on a

$$u = e^{\pm ie^x}$$

et u, de module 1, décrit le cercle de rayon 1 autour de l'origine.

Donc quand u décrit γ, z décrit la double infinité de parallèles à l'axe des x

$$y = 2k\pi \pm \pi/2. \quad (k \text{ entier}).$$

Considérons, en particulier, la parallèle Γ à l'axe des x

$$y = \pi/2,$$

qui est la médiane de la demi-tranche primaire comprise entre les deux lignes A et B, et à laquelle correspond la valeur

$$u = e^{ie^x}$$

Cette valeur est égale à 1 pour toutes les solutions de l'équation $e^x = 2k\pi$ (k positif). Donc la ligne Γ porte une infinité de représentants du point $u = 1$, d'abscisses successives

$$x = \log 2k\pi \quad (k \text{ entier positif})$$

L'écart de ces points consécutifs tend vers zéro quand k augmente indéfiniment.

Voici encore une observation préliminaire. Le module de u a pour expression

$$|u| = e^{e^x \cos y}$$

Il en résulte que u tend toujours vers 1 quand x tend vers l'infini négatif. Quand x tend vers l'infini positif, u tend vers l'infini ou vers zéro selon que $\cos y$ est positif ou négatif. En particulier, dans la tranche entre A et B, la limite est l'infini ou zéro selon que y est $< $ ou $> \pi/2$, c'est-à-dire selon que z varie d'un côté ou de l'autre de la ligne Γ.

Arrivons maintenant au partage des tranches.

Par chaque point $z_1, z_2, \ldots z_n, \ldots$ représentant le point $u = 1$ sur la ligne Γ menons la ligne Δ qui correspond au segment $a + b$ de l'axe réel dans le plan (u) entre les points $u = 0$ et $u = \infty$. En vertu de l'observation précédente, cette ligne (qui ne peut recouper A) se compose de deux branches situées de part et d'autre de Δ, u tendant vers zéro d'un côté, et vers l'infini de l'autre, quand x tend vers l'infini positif.

Les lignes successives $\Delta_1, \Delta_2, \ldots \Delta_n, \ldots$ dont la suite est illimitée sont emboîtées l'une dans l'autre, car elles ne peuvent se rencontrer. Elles partagent la tranche entre A et B en une suite illimitée de tranches subdivisionnaires, lesquelles sont de nouvelles tranches de période, car la ligne $a + b$ correspondant à A et B ne partage par le plan (u). Le procédé de subdivision s'arrête ici.

Ce mode de subdivision s'applique de la même manière à la seconde moitié de la tranche primaire considérée. Elle s'étend à toutes les tranches primaires, car le comportement de la fonction est identique dans toutes ces tranches.

4. Remarque. — De prime abord, les résultats précédents ont une apparence paradoxale. Le nombre des tranches du plan augmente à l'infini dans un domaine qui s'écarte à l'infini dans le plan (z). Or ces tranches sont toujours comprises entre deux droites parallèles, par exemple, les lignes $y = 0$ et $y = \pi$. Leur épaisseur tend vers zéro et elles sont infiniment effilées à la limite, sans cesser de représenter tout le plan (u). Pour l'expliquer, il faut remarquer que quand x tend vers l'infini dans une tranche, la dérivée de u tend vers l'infini, ainsi à un écart infiniment petit dans le plan (z) correspond un écart de grandeur finie dans le plan (u).

5. Prolongement du partage des tranches. — Dans le cas de la fonction $e^{(e^z)}$ considéré au n° 3, la tranche primaire se partage en une suite illimitée de tranches de période et le partage s'arrête là. Mais ce cas n'est pas exclusif. On peut construire des fonctions telles que les tranches nouvelles se partagent à leur tour, et ainsi de suite consécutivement un nombre fini quelconque de fois. En voici la vérification.

Considérons la suite des fonctions, définies de proche en proche,

$$u = e^z, \quad u_1 = e^u, \quad u_2 = e^{u_1}, \ \ldots \ u_n = e^{u_{n-1}}$$

Toutes ces fonctions sont holomorphes de période 2π, admettent les deux valeurs exceptionnelles 0 et ∞ au sens de Picard. Chacune d'elles admet au sens généralisé une valeur exceptionnelle de plus que la précédente (n° 2).

Toutes ces valeurs exceptionnelles sont réelles, mais le passage d'une fonction à la suivante les déplace sur l'axe réel. On constate de proche en proche que la dérivée de ces fonctions ne s'annule pas, ce qui assure l'uniformité de la correspondance des lignes.

Nous avons étudié la fonction u_1 au n° 3. Considérons maintenant le passage de u_1 à u_2. Les valeurs exceptionnelles de u_1 sont 0, 1, ∞ qui deviennent 1, e, ∞ pour u_2. Les lignes du partage en tranches de u_1 correspondent aux segments $0 — \infty$ de l'axe réel du plan u. Ce segment est remplacé par celui de 1 à ∞ pour u_2.

Le partage des tranches de u_1 en sous-tranches pour u_2 se fera donc au moyen des lignes qui correspondent au segment $0 — \infty$ du même axe réel. Ces lignes, infinies dans les deux sens, seront des transversales du plan. Elles ne peuvent couper les lignes de partage de u_1 car les segments de l'axe réel des u ne se coupent pas. Elles divisent bien la tranche considérée de u_1 en parties infinies.

Des considérations analogues s'appliquent au passage d'une autre fonction u à la suivante. Dans tous les cas les lignes de partage des tranches s'écartent à l'infini au fur et à mesure de leur description.

6. La démonstration du théorème de Picard. — Le théorème de Picard affirme qu'une fonction holomorphe ne peut admettre, au sens strict, deux valeurs exceptionnelles finies. J'ai essayé dans l'article précédent de démontrer ce théorème par des considérations topologiques. Cette démonstration est vicieuse, mais nous avons maintenant de sérieuses raisons de croire *qu'une telle démonstration est impossible.*

Supposons, par impossible, que la fonction holomorphe $f(z)$ admette, au sens strict, les deux valeurs exceptionnelles finies 0 et α. On en déduit immédiatement que le partage en d'autres de ses tranches de période peut se poursuivre à l'infini. Mais les lignes de partage doivent de proche en proche reculer à l'infini. C'est dans la méconnaissance de cette loi que consiste l'erreur de la démonstration rappelée ci-dessus (n° 7 de l'art.).

Dans les exemples, étudiés dans les n^{os} précédents, de fonctions douées d'un système de valeurs exceptionnelles généralisées, la succession du partage des tranches ne peut se poursuivre à l'infini, ce qui est une différence essentielle, mais ces exemples ne sont que des cas particuliers. Si, plus généralement, il existe des fonctions douées de valeurs exceptionnelles au sens généralisé dont les tranches soient indéfiniment partageables, leur comportement topologique est identique à celui auquel conduit l'hypothèse d'une fonction douée de trois valeurs exceptionnelles au sens de Picard. Si de telles fonctions existent, le théorème de Picard ne relève pas de la topologie. La question reste ouverte et je souhaite que quelque mathématicien plus jeune que moi réussisse à la trancher.

Annales de la Société Scientifique de Bruxelles, Tome **70**, n° II, p. 87-95

Produits scalaires à annulation symétrique

PAR

Robert BALLIEU

RÉSUMÉ. — Tout produit scalaire sur un espace vectoriel, par rapport à un antiautomorphisme w du corps de base, s'annulant symétriquement, donne lieu à une relation d'orthogonalité, jouissant de propriétés, classiques dans les cas particuliers des produits scalaires symétriques, antisymétriques, hermitiens et hermitiens gauches.

Parmi les produits scalaires s'annulant symétriquement, il en est d'exceptionnels, s'annulant sur un sous-espace de codimension 1, orthogonal à ses supplémentaires. Les autres sont caractérisés par l'existence d'une constante d telle que l'on ait, pour des vecteurs quelconques ξ et η,

$$w(f(\xi, \eta) = df(\eta, \xi). \tag{1}$$

Dans ce cas, on a, pour tout scalaire x,

$$w^2(x) = dxd^{-1}; \tag{2}$$

de plus,

$$dw(d) = 1. \tag{3}$$

Inversement, étant donné un antiautomorphisme w d'un corps **K** et un scalaire d, satisfaisant à (2) et à (3), on peut définir, sur tout espace vectoriel sur le corps **K**, un produit scalaire, satisfaisant à (1). Si la dimension de l'espace est > 1, on peut définir un tel produit non exceptionnel.

On indique des exemples non classiques de produits scalaires, s'annulant symétriquement.

1. DÉFINITIONS ET NOTATIONS. — Dans cet article, **K** désigne un corps; un corps commutatif est appelé « *champ* »; w est un antiautomorphisme de **K**, c'est-à-dire un automorphisme de **K** sur son opposé; les lettres a, b, c, d, u, x, y, avec indice ou non, représentent des scalaires, c'est-à-dire des éléments de **K**.

L désigne un espace vectoriel à gauche sur **K**; M, P, R, avec indice ou non, sont des sous-espaces (vectoriels) de L; ξ, ζ, η, avec indice ou non, sont des vecteurs.

Un *produit scalaire par rapport à* w est une application f de $L \times L$ dans **K**, telle que

(4) $$f(a_1\xi_1 + a_2\xi_2, b_1\eta_1 + b_2\eta_2) = \sum_{i,j} a_i f(\xi_i, \eta_j) w(b_j), (1).$$

Ce produit est à *annulation symétrique* ou *s'annule symétriquement* si $f(\xi, \eta) = 0$ entraîne $f(\eta, \xi) = 0$.

Une *prérelation d'orthogonalité* sur L est une relation binaire, symétrique, linéaire, vérifiée par le vecteur nul et un vecteur quelconque. La linéarité de la relation signifie que la relation est satisfaite par ξ et $a\eta + b\zeta$, si elle est satisfaite par ξ et chacun des vecteurs η et ζ. Deux parties de L sont *orthogonales* si tout vecteur de l'une et tout vecteur de l'autre vérifient la relation. Un vecteur *isotrope* est un vecteur orthogonal à soi-même. Le *complément orthogonal* d'un sous-espace P dans le sous-espace M est l'ensemble des vecteurs de M orthogonaux à P; c'est un sous-espace de M. Le sous-espace M est *régulier* si le complément orthogonal de M dans soi-même est réduit au vecteur nul. Une famille de vecteurs est *orthogonale* si deux vecteurs d'indices différents sont orthogonaux.

Une *relation d'orthogonalité* sur L est une prérelation d'orthogonalité telle que la dimension du complément orthogonal, dans un M de dimension finie m, d'un sous-espace P de M, de dimension p, est $\geqslant (m - p)$.

2. Soient L, avec une prérelation d'orthogonalité, et le complément orthogonal L_0 de L dans soi-même. Tout supplément M de L_0 dans L, c'est-à-dire tout sous-espace M tel que L soit somme directe de L_0 et de M, est régulier; la somme d'un sous-espace régulier et de son complément orthogonal est directe; tout sous-espace régulier est compris dans un supplément de L_0 dans L.

Si un vecteur est combinaison linéaire, à coefficients tous non nuls, d'une famille (finie), libre et orthogonale, de vecteurs non isotropes, ce vecteur n'est orthogonal à aucun élément de la famille. En particulier, si L admet une base orthogonale de vecteurs non isotropes, L est régulier.

---

(1) On sait qu'on ramène à ce cas les applications f de L × L dans **K**, satisfaisant à

$$f(a_1\xi_1 + a_2\xi_2, b_1\eta_1 + b_2\eta_2) = \sum_{i,j} w_1(a_i) f(\xi_i, \eta_j) w(b_j),$$

où w_1 est un automorphisme de **K**. On sait aussi que si la relation précédente a lieu pour une application non nulle f de L × L dans **K** et deux applications w_1 et w de **K** sur soi-même, w_1 est un automorphisme et w un antiautomorphisme de **K**.

Toute famille orthogonale de vecteurs non isotropes est libre.

Tout sous-espace admettant une base orthogonale de vecteurs non isotropes est compris dans un sous-espace ayant la même propriété et dont le complément orthogonal dans L est formé de vecteurs isotropes.

3. Soit L avec une relation d'orthogonalité. Le complément orthogonal dans L d'un sous-espace propre, de dimension finie, ne se réduit pas au vecteur nul. En particulier, si la dimension de L est strictement supérieure à 1, tout vecteur est orthogonal à un vecteur non nul. Le complément orthogonal d'un sous-espace propre de dimension infinie peut se réduire au vecteur nul.

Quel que soit le sous-espace régulier P de dimension finie, L est somme directe de P et du complément orthogonal R de P dans L; si L est de plus régulier, R est aussi régulier.

Si la codimension dans L du complément orthogonal L_0 de L dans soi-même est finie, L est somme directe de L_0 et de deux sous-espaces réguliers L_1 et L_2, jouissant des propriétés suivantes :

1° L_1 et L_2 sont orthogonaux, (1);

2° L_1 admet une base orthogonale de vecteurs non isotropes;

3° L_2 est formé de vecteurs isotropes.

Le sous-espace $(L_0 + L_2)$, supplémentaire de L_1 dans L, est orthogonal à L_1 et est formé de vecteurs isotropes.

De plus, si (ξ_i), $(i \in I)$, est une famille orthogonale de vecteurs non isotropes, il existe une telle décomposition où L_1 admet une base orthogonale de vecteurs non isotropes, contenant la famille (ξ_i), qui est nécessairement finie.

La dimension d'un espace vectoriel avec relation d'orthogonalité, régulier et formé de vecteurs isotropes, est paire ou infinie. Aucun élément d'une base d'un tel espace n'est orthogonal à chacun des autres éléments de cette base.

4. Soit un espace vectoriel L, muni d'un produit scalaire f par rapport à w, s'annulant symétriquement. La relation $f(\xi,\eta) = 0$ entre deux vecteurs ξ et η, est une relation d'orthogonalité. En effet, elle est symétrique et linéaire; de plus, pour tout η, on a $f(0, \eta) = 0$. Enfin, soit un sous-espace M de dimension finie et un sous-espace P de M. On considère une base $(\xi_1, \ldots, \xi_m)$ de

---

(1) L_1 et L_2 sont aussi orthogonaux à L_0.

M telle que $(\xi_1, \ldots, \xi_p)$ soit une base de P. Si $a_{i,j} = f(\xi_i, \xi_j)$, on a

$$f(\Sigma x^i \xi_i, \Sigma y^i \xi_i) = \sum_{i,j} x^i a_{i,j} w(y^j).$$

Le complément orthogonal R de P dans M est l'ensemble des vecteurs $\Sigma x^i \xi_i$ tels que

$$\sum_i x^i a_{i,j} = 0, \quad (j = 1, 2, \ldots, p). \tag{5}$$

R est isomorphe à l'espace des solutions $(x^1, \ldots, x^m)$ de (5); sa dimension est $\geqslant (m - p)$.

Les propriétés indiquées aux nos 2 et 3 s'appliquent donc à L.

La régularité de M par rapport à l'orthogonalité ainsi définie revient à la régularité de M par rapport au produit scalaire f, c'est-à-dire au fait que le vecteur nul est le seul vecteur ξ de M, tel que $f(\xi, \eta)$ soit nul pour tout vecteur η de M.

5. Soit l'espace vectoriel L du n° 4.

Lemme I. — Soit un sous-espace M *de* L, *admettant une base orthogonale* (ξ_i), $(i \in I)$, *de vecteurs non isotropes; on pose* $a_i = f(\xi_i, \xi_i)$. *Alors, les produits* $w(a_i)\, a_i^{-1}$ *sont égaux entre eux; si* d *est leur valeur commune, on a, sur* M,

$$w(f(\xi, \eta)) = df(\eta, \xi). \tag{6}$$

On considère deux éléments distincts ξ_1 et ξ_2, appartenant à la base.

On a

$$f(x_1\xi_1 + x_2\xi_2, y_1\xi_1 + y_2\xi_2) = x_1 a_1 w(y_1) + x_2 a_2 w(y_2). \tag{7}$$

En vertu de l'annulation symétrique de f et de l'antiautomorphisme w, les relations

$$x_1 a_1 y_1 + x_2 a_2 y_2 = 0, \tag{8}$$

$$w^{-1}(y_1)\, a_1 w(x_1) + w^{-1}(y_2)\, a_2 w(x_2) = 0, \tag{9}$$

$$w^2(x_1)\, w(a_1)\, y_1 + w^2(x_2)\, w(a_2)\, y_2 = 0, \tag{10}$$

sont vérifiées simultanément par des éléments x_1, x_2, y_1, y_2 de **K**. Donc, pour x_1 et x_2 quelconques dans **K**, il existe dans **K** un élément $d(x_1, x_2)$, non nul, tel que l'on ait

$$w^2(x_1)\, w(a_1) = d(x_1, x_2)\, x_1\, a_1, \tag{11}$$

$$w^2(x_2)\, w(a_2) = d(x_1, x_2)\, x_2\, a_2; \tag{12}$$

si $x_1 = x_2 = 0$, $d(x_1, x_2)$ peut être pris arbitrairement dans **K**. En vertu de (11), $d(x_1, x_2)$ ne dépend pas de x_2 si $x_1 \neq 0$; de même, $d(x_1, x_2)$ ne dépend pas de x_1 si $x_2 \neq 0$. Donc, les valeurs $d(x_1, x_2)$ sont toutes égales entre elles, lorsque $(x_1, x_2) \neq (0,0)$. En désignant par d leur valeur commune, non nulle, on a

$$d = w(a_1)\, a_1^{-1} = w(a_2)\, a_2^{-1}, \tag{13}$$

$$w^2(x)\, w(a_1) = dxa_1, \tag{14}$$

$$w^2(x) = dxd^{-1}. \tag{15}$$

Si on prend pour ξ_1 un vecteur fixe de la base (ξ_i) et pour ξ_2 un vecteur quelconque de cette base, la relation (13) montre que les produits $w(a_i)\, a_i^{-1}$ sont bien tous égaux à d.

On a, pour des vecteurs $\xi = \Sigma\, x_i \xi_i$ et $\eta = \Sigma\, y_i \xi_i$ de M,

$$f(\xi, \eta) = \Sigma\, x_i\, a_i\, w(y_i),\ f(\eta, \xi) = \Sigma\, y_i\, a_i\, w(x_i),$$

$$w(f(\xi, \eta)) = \Sigma\, w^2(y_i)\, w(a_i)\, w(x_i) = \Sigma\, dy_i d^{-1}\, w(a_i)\, w(x_i) = df(\eta, \xi).$$

Lemme II. — S'il existe un vecteur ξ non isotrope, orthogonal à un sous-espace P *de vecteurs isotropes, sur lequel f ne se réduit pas au produit nul, le corps* **K** *est un champ de caractéristique deux; w est l'automorphisme identique; sur* P, *le produit scalaire est symétrique.*

En effet, soit un espace vectoriel sur un corps **K**, muni d'un produit scalaire par rapport à w, non nul et pour lequel tout vecteur soit isotrope. Alors, (1), **K** est un champ; w est l'identité; le produit scalaire est antisymétrique.

Il reste à prouver que la caractéristique est deux. Pour cela, soient deux vecteurs η et ζ de P tels que

$$u = f(\eta, \zeta) = -f(\zeta, \eta) \neq 0.$$

On a, compte tenu de l'annulation symétrique, en désignant par a le scalaire non nul $f(\xi, \xi)$,

$$f(\xi - au^{-1}\eta, \xi + \zeta) = f(\xi, \xi) - au^{-1} f(\eta, \zeta) = a - a = 0,$$

$$0 = f(\xi + \zeta, \xi - au^{-1}\eta) = a - au^{-1} f(\zeta, \eta) = 2a.$$

Remarque. — Les deux vecteurs $(\xi - au^{-1}\eta)$ et $(\xi + \zeta)$ sont orthogonaux entre eux et non isotropes.

---

(1) Cfr. JACOBSON, N., *Lectures in abstract algebra*, t. II, p. 160, (New-York, Van Nostrand, (1953)).

6. *Théorème I. — Soit un espace vectoriel* L, *muni d'un produit scalaire par rapport à* w, *s'annulant symétriquement. Ou bien le produit scalaire est nul sur un sous-espace de codimension* 1, *orthogonal à l'un de ses supplémentaires dans* L, [1]. *Ou bien, il existe un scalaire* d, *tel que l'on ait, pour des vecteurs* ξ *et* η *quelconques,*

$$w(f(\xi, \eta)) = df(\eta, \xi). \tag{16}$$

S'il existe deux vecteurs ξ_1 et ξ_2, orthogonaux entre eux et non isotropes, on pose

$$a_1 = f(\xi_1, \xi_1), \; d = w(a_1)\, a_1^{-1}.$$

On considère l'espace vectoriel R, engendré par les vecteurs ξ_1 et ξ_2 et par deux vecteurs η et ζ quelconques. En vertu du n° 3, R est somme directe d'un sous-espace M, admettant une base orthogonale de vecteurs non isotropes, comprenant ξ_1 et ξ_2, et d'un sous-espace P, orthogonal à M et formé de vecteurs isotropes.

Sur M, on a, avec le scalaire d ci-dessus, la propriété de l'énoncé, (lemme I). Cette propriété s'applique aussi à P. Si, sur P, le produit scalaire f se réduit au produit nul, le fait est trivial. Sinon, (lemme II), **K** est un champ de caractéristique deux; w est l'identité; d est égal à 1; sur P, f est symétrique.

La propriété de l'énoncé s'applique aux vecteurs η et ζ. Car, on a

$$\eta = \eta' + \eta'', \; \zeta = \zeta' + \zeta'',$$

où $\eta', \zeta' \in \mathrm{M}$, $\eta'', \zeta'' \in \mathrm{P}$. De là,

$$f(\eta, \zeta) = f(\eta', \zeta') + f(\eta'', \zeta''), \; f(\zeta, \eta) = f(\zeta', \eta') + f(\zeta'', \eta''),$$

$$w(f(\eta, \zeta)) = w(f(\eta', \zeta')) + w(f(\eta'', \zeta''))$$

$$= df(\zeta', \eta') + df(\zeta'', \eta'') = df(\zeta, \eta).$$

Le théorème est démontré dans ce cas, puisque η et ζ sont deux vecteurs quelconques.

Il reste à envisager le cas où il n'existe pas deux vecteurs non isotropes, orthogonaux entre eux, et où le cas exceptionnel de l'énoncé n'est pas réalisé. Ou bien tous les vecteurs sont isotropes; alors, le théorème résulte de la propriété citée au lemme II. Ou bien, il existe un vecteur non isotrope ξ; L est somme directe de $\mathrm{M} = \mathbf{K}\xi$ et du complément orthogonal P de M dans L; P est

---

[1] Il revient au même de dire que le complément orthogonal de L en soi-même est de codimension 0 ou 1 par rapport à L.

formé de vecteurs isotropes; f ne se réduit pas sur P au produit nul; alors, (lemme II), la propriété de l'énoncé s'applique à M et à P, en prenant $d = 1$; elle est valable pour L, avec la même valeur de d.

Remarque. — Dans l'éventualité envisagée en dernier lieu, il existe deux vecteurs non isotropes, orthogonaux entre eux, (lemme II, rem.). Donc en dehors du cas exceptionnel, (qui comprend les espaces vectoriels de dimension 1), et du cas d'un produit scalaire antisymétrique, il existe toujours deux vecteurs non isotropes, orthogonaux entre eux.

Scolie. — S'il existe un sous-espace, formé de vecteurs isotropes, sur lequel f ne se réduit pas au produit nul, f est antisymétrique au sens large, (1).

En effet, (lemme II), **K** est alors un champ et w est l'identité sur **K**. Sur le sous-espace, f est antisymétrique et non nul; donc, $d = -1$; (16) s'écrit $f(\xi, \eta) = -f(\eta, \xi)$. En particulier, si f n'est pas antisymétrique au sens large, il n'existe pas de sous-espaces réguliers non nuls, formés de vecteurs isotropes. De là, (n° 3),

Corollaire. — Soit un espace vectoriel, de dimension finie sur un corps **K**, *muni d'un produit scalaire par rapport à un antiautomorphisme w de* **K**, *non antisymétrique au sens large et s'annulant symétriquement. Alors, l'espace admet une base orthogonale. S'il est régulier, il admet une base orthogonale de vecteurs non isotropes; toute famille orthogonale de vecteurs non isotropes peut être complétée en une telle base.*

Théorème II. — Soit un espace vectoriel, muni d'un produit scalaire f par rapport à w, pour lequel il existe un scalaire d tel que l'on ait

$$w(f(\xi, \eta)) = df(\eta, \xi), \tag{17}$$

pour des vecteurs ξ et η quelconques. Alors, f s'annule symétriquement. Pour tout vecteur non isotrope ξ, on a, posant $a = f(\xi, \xi)$,

$$d = w(a)\, a^{-1}. \tag{18}$$

---

(1) On veut dire que $f(\xi, \eta) = -f(\eta, \xi)$, sans exiger, (si la caractéristique est deux), que $f(\xi, \xi) = 0$.

Si f n'est pas le produit nul, on a

(19) $$w^2(x) = dxd^{-1}$$

pour un scalaire x quelconque; de plus,

(20) $$dw\,(d) = 1.$$

L'annulation symétrique est évidente. La relation (18) est immédiate. On applique w aux deux membres de (17); il vient

(21) $$w^2\,(f\,(\xi, \eta)) = w\,(f\,(\eta, \xi))\,w\,(d) = df\,(\xi, \eta)\,w\,(d).$$

Si f n'est pas le produit scalaire nul, f est une application de $\mathrm{L} \times \mathrm{L}$ sur $\mathbf{K}$. On a donc, pour tout x,

(22) $$w^2\,(x) = dxw\,(d).$$

De là, pour $x = 1$, la formule (20), puis la formule (19).

L'antiautomorphisme w et la constante d ne sont soumis qu'aux seules conditions (19) et (20).

Soient en effet un espace vectoriel L sur un corps $\mathbf{K}$, (de dimension >1), une base (ξ_i), $(i \in \mathrm{I})$, de L, un antiautomorphisme w de $\mathbf{K}$ pour lequel il existe une constante d telle que (19) et (20) aient lieu.

Etant donnés deux indices distincts i et j, on prend deux éléments $a_{i,j}$ et $a_{j,i}$ de $\mathbf{K}$, liés par

(23) $$w\,(a_{i,j}) = da_{j,i};$$

(on peut choisir arbitrairement $a_{i,j}$ et définir $a_{j,i}$ par la relation (23)); on a alors

$$w^2\,(a_{i,j}) = w\,(a_{j,i})\,w\,(d),\ da_{i,j}\,d^{-1} = w\,(a_{j,i})\,d^{-1},$$

$$w\,(a_{j,i}) = da_{i,j}.$$

On prend aussi des éléments $a_{i,i}$ tels que

$$w\,(a_{i,i}) = da_{i,i};$$

on peut en tout cas poser $a_{i,i} = 0$. Cela étant, on dispose d'éléments $a_{i,j}$, $((i, j) \in \mathrm{I} \times \mathrm{I})$, satisfaisant à (23).

Soient les vecteurs $\xi = \Sigma\, x_i\xi_i$ et $\eta = \Sigma\, y_i\xi_i$. La fonction f, définie par

$$f\,(\xi, \eta) = \sum_{i,j} x_i a_{i,j} w\,(y_j),$$

est un produit scalaire sur L par rapport à w. De plus, puisque

$$w\,(f\,(\xi, \eta)) = \sum_{i,j} w^2\,(y_j)\,w\,(a_{i,j})\,w\,(x_i) = df\,(\eta, \xi),$$

f satisfait aux hypothèses du théorème II. Il suffit qu'un des $a_{i,j}$, $(i \neq j)$, ne soit pas nul pour que f ne soit pas le produit nul; de plus, en prenant alors tous les $a_{i,i} = 0$, f n'est pas un produit exceptionnel.

7. Exemples. — 1° On prend le champ **C** des nombres complexes et, pour w, l'application faisant correspondre à x son conjugué $\bar{x}$; d est un nombre quelconque de module 1. Les solutions de $w(x) = dx$ sont les produits d'une racine carrée de d^{-1} par un nombre réel.

2° On prend le corps des quaternions réels et, pour w, l'anti-automorphisme défini par

$$w(x) = a\bar{x}a^{-1},$$

si a est un quaternion non nul fixe et si $\bar{x}$ désigne le quaternion conjugué de x; la constante $d = a\bar{a}^{-1}$, quotient de a^2 par la norme de a, convient ainsi que la constante $d = -a\bar{a}^{-1}$; pour la première constante, les solutions de $w(x) = dx$ sont les multiples réels de $\bar{a}$; pour la seconde, ces solutions sont les produits de $\bar{a}$ par les quaternions-vecteurs, (c'est-à-dire à partie réelle nulle).

8. Remarques. — 1° Les conditions (19) et (20) sont indépendantes. Par exemple, si w est l'automorphisme identique sur un champ **K**, (19) est vérifiée pour toute constante d non nulle; (20) ne l'est que pour $d = 1$ ou -1. Dans le second exemple du n° 7, il existe des constantes d satisfaisant à (20) et ne satisfaisant pas à (19).

2° La condition (19) pour un antiautomorphisme w d'un corps **K** entraîne que l'élément $dw(d)$ est invariant pour w et appartient au centre de **K**.

3° Les exemples du n° 7 montrent que, w étant donné, les conditions (19) et (20) ne déterminent pas univoquement d. S'il y a une constante d_0 satisfaisant à (19) et (20), les autres solutions d sont les produits de d_0 par les éléments c du centre de **K**, tels que

$$c\,w(c) = 1.$$

Annales de la Société Scientifique de Bruxelles, Tome **70**, n° II, p. 96-116

A propos de la notion de fonction aléatoire

PAR

H. BRENY
Dr. Sc.

1. Introduction.

Comme on le sait, la notion de fonction qui, en soi, est une notion tout à fait abstraite et générale, s'est en fait introduite en mathématiques par son analogue concret de « fonction du temps ». Cela est bien clair, p. ex., dans les écrits de Newton sur la « méthode des fluxions », et même chez les auteurs qui, vers 1600, s'occupent de problèmes géométriques comme la recherche des tangentes, la notion de fonction ne se superpose à celle de courbe que par un intermédiaire cinématique. Ce n'est en somme que récemment que la notion abstraite, statique, ensembliste, de fonction est devenue d'un usage plus ou moins général. Ce qui s'est produit pour les fonctions de l'analyse s'est aussi produit pour les fonctions aléatoires. Sans doute a-t-on très vite vu, du moins dans l'abstrait, qu'une fonction aléatoire n'est pas nécessairement fonction du temps; mais, concrètement, c'est bien dans le concept de variabilité temporelle que s'est tout d'abord réalisée la notion de fonction aléatoire.

La source principale de ce mouvement d'idées a sans doute été l'étude des phénomènes économiques ; il est significatif de constater, p. ex., que, dans le grand traité de statistique de Kendall, les chapitres relatifs aux fonctions aléatoires sont illustrés d'exemples presqu'exclusivement économiques (notamment la fameuse série de prix établie par Beveridge) ; c'est alors qu'à pris naissance le terme même de « série chronologique » (« time-series ») qui, même à l'heure actuelle, est parfois encore employé comme synonyme de « fonction aléatoire », ce qui met très clairement en lumière la thèse énoncée tantôt. Une autre source, sans doute non moins

importante, doit être cherchée dans les travaux d'un certain nombre de techniciens de l'électronique qui, vu la précision toujours croissante de leurs mesures, ont inévitablement rencontré des phénomènes, comme le bruit de fond, dont le caractère de « fluctuations aléatoires » s'est très vite imposé à leur attention. On pourrait penser que la mécanique statistique, qui est une science relativement ancienne, aurait contribué à éclaircir la notion même de fonction aléatoire car elle en offre, comme nous le verrons, des exemples particulièrement clairs et importants ; mais il ne semble pas que cela se soit produit.

2. Premier point de vue : définition par la loi temporelle.

2,1. Imaginons donc un économiste ou un électronicien, quelque peu au courant des méthodes de l'analyse statistique, confronté avec une grandeur qui, au cours du temps, fluctue d'une manière régie, au moins partiellement, par le hasard. Le physicien n'aura sans doute aucune peine à imaginer qu'il répète son expérience, donnant ainsi lieu à une population de fonctions dont chaque élément est une réalisation particulière du phénomène aléatoire qu'il étudie ; pour l'économiste, une telle idée est un tant soit peu plus abstraite. Pour l'un comme pour l'autre, cependant, la première idée qui vient à l'esprit est de considérer, en fonction du paramètre t, la variation de la loi de probabilité de la valeur à l'instant t, $\mathbf{x}(t)$ [1], de la fonction aléatoire $\mathbf{x}(.)$ considérée. Il n'y a là, somme toute, aucune idée nouvelle : pour t fixé, $\mathbf{x}(t)$ est une variable aléatoire au sens ordinaire, dont nous désignerons par $\mathrm{F}_1(\xi ; t)$ la fonction de répartition :

$$\Pr[\mathbf{x}(t) \leqslant \xi] = \mathrm{F}_1(\xi ; t).$$

Pour un ensemble de valeurs de t, on a donc un ensemble de variables aléatoires, donc un ensemble de fonctions de répartition, ce qui se marque analytiquement par le fait que $\mathrm{F}(\xi ; t)$ ne dépend pas seulement de la variable indicatrice ξ mais encore du paramètre t. Mais on sait que (au contraire des ensembles « nus » de la théorie des ensembles, mais analogues en cela aux ensembles « qualifiés » des autres doctrines mathématiques) les ensembles

(1) Les grandeurs aléatoires sont notées par des symboles en caractères gras et distinguées ainsi des valeurs qu'elles prennent pour une épreuve déterminée, valeurs qui sont notées par les mêmes symboles en caractères ordinaires.

de variables aléatoires ne sont pas entièrement définis par la simple donnée de leurs éléments : il faut encore définir les relations mutuelles de ceux-ci. De ce point de vue, l'élément qui s'offre le plus immédiatement à la considération est la fonction de répartition conjointe de deux valeurs de la fonction :

$$F_1(\xi_1, \xi_2; t_1, t_2) \equiv \Pr[\mathbf{x}(t_1) \leqslant \xi_1, \mathbf{x}(t_2) \leqslant \xi_2].$$

On peut évidemment continuer dans ce sens, mais, historiquement, on peut bien dire que l'on s'en est très souvent tenu à cette « étude du 2e ordre » des fonctions aléatoires. On y voit immédiatement apparaître les éléments suivants :

la moyenne locale : $m(t) = \mathcal{E}\,\mathbf{x}(t) = \int_{-\infty}^{\infty} \xi\, d\, F_1(\xi; t)$

la variance locale : $\sigma^2(t) = \mathcal{E}\,[\mathbf{x}(t) - m(t)]^2 = \int_{-\infty}^{\infty} [\xi - m(t)]^2\, dF_1(\xi; t)$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [\xi_1 - m(t)]^2\, d\, F_2(\xi_1, \xi_2; t, t_2)$$

l'autocovariance : $c(t, \delta) = \mathcal{E}\,[\mathbf{x}(t) - m(t)]\,[\mathbf{x}(t+\delta) - m(t+\delta)]$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [\xi_1 - m(t)]\,[\xi_2 - m(t+\delta)]\, d\, F_2(\xi_1, \xi_2; t, t+\delta)$$

et l'une des premières remarques que l'on a faites à ce sujet est celle-ci : on n'a pas nécessairement : $\lim_{\delta \to 0} c(t, \delta) = \sigma^2(t)$. Si cette relation a lieu pour la valeur t_0 de t, on dit que la fonction aléatoire $\mathbf{x}(.)$ est continue en moyenne quadratique pour $t = t_0$.

L'étude des fonctions pour lesquelles $m(t)$, $\sigma^2(t)$ et $c(t, \delta)$ ne dépendent pas de t — fonctions qui sont dites « stationnaires au 2d ordre » — a donné lieu à un grand nombre de travaux.

2,2. Si l'on considère un nombre quelconque de valeurs de $t : t_1, ..., t_n$, on peut former l'expression :

$$F_n(\xi_1, ..., \xi_n; t_1, ..., t_n) \equiv \Pr[\mathbf{x}(t_1) \leqslant \xi_1, ..., \mathbf{x}(t_n) \leqslant \xi_n]. \qquad (1)$$

Supposons que ces fonctions de répartition soient définies pour toutes les valeurs possibles de n et des t_i, en accord avec les conditions évidemment nécessaires de compatibilité et de symétrie, du type :

$$F_{n+m}(\xi_1, ..., \xi_n, \infty, ..., \infty\ ; t_1, ..., t_{n+m}) = F_n(\xi_1, ..., \xi_n\ ; t_1, ..., t_n) \quad (2)$$

$$F_n(P\xi_1, ..., P\xi_n; Pt_1, ..., Pt_n) = F_n(\xi_1, ..., \xi_n; t_1, ..., t_n) \qquad (3)$$

P désignant une permutation quelconque du groupe symétrique d'ordre $n!$. Nous allons voir que ces conditions nécessaires sont aussi, en un sens, suffisantes, de sorte que, moyennant (2) et (3), (1) définit entièrement ce que l'on appelle la « loi temporelle » de la fonction aléatoire $\mathbf{x}(.)$; il reste toutefois à préciser dans quelle mesure on peut dire que cette fonction aléatoire est définie par sa loi temporelle.

2,21. Il nous faut, pour le voir clairement, rappeler très brièvement quelques notions fondamentales du calcul des probabilités. On sait qu'un « espace probabilisé » (ou « catégorie d'épreuves ») est un ensemble E, de nature quelconque, dans lequel on a défini deux choses : d'une part, une famille $\mathcal{F}$ de sous-ensembles de E jouissant des trois propriétés suivantes :

A) $E \in \mathcal{F}$,

B) si $F \in \mathcal{F}$, il en est de même du complémentaire de F par rapport à E : $(E - F) \in \mathcal{F}$,

C) si $F_i \in \mathcal{F}$ pour $i = 1, 2, \ldots, n, \ldots$, il en est de même de $\mathbf{U}_{i=1}^{\infty} F_i$, (une telle famille est appelée un « corps borélien »); d'autre part, une fonction d'ensemble, Pr(F), définie pour tout élément de la famille $\mathcal{F}$, réelle, comprise entre 0 et 1 et telle que

a) $\Pr(E) = 1$

b) si $F_i \in \mathcal{F}$ et $F_i \cap F_j$ est vide pour tout i et tout j, $\Pr[\mathbf{U}_{i=1}^{\infty} F_i] = \Sigma_{i=1}^{\infty} \Pr(F_i)$ (additivité complète).

Dans l'ensemble des fonctions réelles d'une variable réelle, $\chi \equiv \{x(.)\}$, considérons alors les sous-ensembles définis par des conditions de la forme

$$a'_1 < x(t_1) \leqslant a''_1, \ldots, a'_n < x(t_n) \leqslant a''_n,$$

les t_i, a'_i et a''_i étant des nombres réels quelconques et n un entier > 0 quelconque. La famille $\mathcal{F}_0$ de tous ces sous-ensembles de χ n'est pas un corps borélien, mais il existe un « corps borélien minimal » engendré par $\mathcal{F}_0$, $\mathcal{F}_T$ (qui comprend $\mathcal{F}_0$ et dont on ne peut rien retrancher sans qu'il cesse soit d'être un corps borélien soit de comprendre $\mathcal{F}_0$). La donnée de la loi temporelle définit sur $\mathcal{F}_0$ une fonction additive d'ensemble, P_1, qui est réelle et comprise

entre 0 et 1; ainsi, on a :

$$P_1[a'_1 < \mathbf{x}(t_1) \leqslant a''_1, \ldots, a'_n < \mathbf{x}(t_n) \leqslant a''_n] = F_n(a''_1, \ldots, a''_n; t_1, \ldots, t_n) - F_n(a'_1 \ldots, a'_n; t_1, \ldots, t_n).$$

Il résulte alors de théorèmes classiques de la théorie de la mesure que l'on peut prolonger P_1, *d'une et une seule manière*, en une fonction P_2 définie sur $\mathcal{F}_T$, y jouissant des propriétés a) et b) ci-dessus et qui, sur $\mathcal{F}_0$, est identique à P_1. On peut alors élargir quelque peu le corps $\mathcal{F}_T$ en un corps $\mathcal{F}'_T$ en y adjoignant tous les sous-ensembles d'ensembles de $\mathcal{F}_T$ de mesure P_2 nulle :

$F \in \mathcal{F}'_T$ s'il existe 3 ensembles F_1, F_2, F_3 tels que :

$$F_1 \in \mathcal{F}_T,\ F_2 \subset F_3,\ F_3 \in \mathcal{F}_T,\ F = F_1 \cup F_2,\ P_2(F_3) = 0;$$

alors, si on définit la fonction d'ensemble P_3 en posant :

$$P_3(F) = P_2(F_1),$$

le corps borélien $\mathcal{F}'_T$ et la mesure P_3 sont encore univoquement déterminés par la loi temporelle.

2,221. Ainsi donc, on peut énoncer que :

« la donnée de la loi temporelle (1) de la fonction aléatoire $\mathbf{x}(.)$ permet de transformer l'ensemble $\mathcal{X}$ des fonctions de t en un espace probabilisé $(\mathcal{X}, \mathcal{F}'_T, P_3)$ où les relations (1) sont toutes vérifiées »;

c'est en ce sens que l'on peut dire que les conditions (2) et (3) sont suffisantes (il n'en serait pas nécessairement ainsi pour des fonctions ne prenant pas leurs valeurs dans un espace euclidien). Malheureusement, il existe une foule de sous-ensembles de $\mathcal{X}$ dont la considération est des plus intéressante et qui n'appartiennent pas à $\mathcal{F}'_T$, qui, par conséquent, n'ont pas de probabilité définie par les raisonnements que je viens de rappeler. Ainsi, si l'on considère un intervalle $\mathcal{J}$ quelconque, le sous-ensemble $\overline{\mathcal{X}}_{\mathcal{J},a}$ de $\mathcal{X}$ formé par les fonctions dont la borne inférieure sur $\mathcal{J}$ est au plus égale à a :

$$\overline{\mathcal{X}}_{\mathcal{J},a} \equiv \{x(.) : \inf_{\mathcal{J}} x(t) \leqslant a\}$$

n'est pas nécessairement P_3-mesurable; s'il ne l'est pas, le nombre : $\inf_{\mathcal{J}} \mathbf{x}(t)$ n'est *pas* une variable aléatoire au sens ordinaire du terme, puisque la probabilité $P_3 [\inf_{\mathcal{J}} \mathbf{x}(t) \leqslant a]$ n'existe pas; conclusion assurément désagréable ! De même, le sous-ensemble $\overline{\overline{\mathcal{X}_{\mathcal{J}}}}$ des fonctions continues sur $\mathcal{J}$ peut ne pas être mesurable, ce qui risque d'entraîner bien des difficultés dans l'application à $\mathbf{x}(.)$ des théorèmes classiques de l'analyse.

2,222. Afin de montrer clairement combien peu une fonction aléatoire est définie par sa loi temporelle, je vais définir d'une manière directe une fonction aléatoire $\mathbf{x}(.)$, de la manière suivante : je considère une variable aléatoire $\boldsymbol{\omega}$ équipartie sur l'intervalle $\Omega \equiv (0,1)$, et un sous-ensemble M de cet intervalle; à chaque valeur ω de $\boldsymbol{\omega}$ j'attache la fonction $x_{\boldsymbol{\omega}}(.)$ définie de la manière suivante :

1°) $x_\omega(t) = 0$ pour tout t si $\omega \in M$

2°) $x_\omega(t) = \begin{cases} 0 \text{ si } \omega \in M \text{ et } t \neq w \\ 1 \text{ si } \omega \in M \text{ et } t = \omega. \end{cases}$

La fonction $x_{\boldsymbol{\omega}}(.)$, dépendant d'une variable aléatoire, est bien une fonction aléatoire.

Je dis que, quels que soient les points en nombre fini $t_1, \ldots, t_n$, on a :

$$\mathcal{P}_1 \equiv \Pr [x_{\boldsymbol{\omega}}(t_1) = 0, \ldots, x_{\boldsymbol{\omega}}(t_n) = 0] = 1 \qquad (4)$$

En effet, pour que, t_i étant donné, $x_\omega(t_i) = 1$, il faut évidemment que $\omega = t_i$ (mais cela ne suffit pas); on a donc :

$$\Pr [x_{\boldsymbol{\omega}}(t_i) = 1] \leqslant \text{mes } \{t_i\} = 0$$

d'où $\Pr [x_{\boldsymbol{\omega}}(t_i) = 0] = 1$ pour tout t_i, d'où la relation (4) (qui reste même valable pour une infinité dénombrable de t_i). Mais, d'autre part, pour que $x_\omega(t) = 0$ pour toute valeur de t, il est nécessaire et suffisant que $\omega \in M$; donc

$$\mathcal{P}_2 \equiv \Pr [x_{\boldsymbol{\omega}}(t) = 0 \text{ pour tout } t]$$

existe si M est mesurable, et vaut alors mes M. On voit donc que l'on peut avoir, simultanément :

ou bien : $\mathcal{P}_1 = 1$, $\mathcal{P}_2$ n'existe pas (M non mesurable)
ou bien : $\mathcal{P}_1 = 1$, $\mathcal{P}_2 = 1$ (mes M $=$ 1; p. ex. M $\equiv \Omega$)
ou bien : $\mathcal{P}_1 = 1$, $0 < \mathcal{P}_2 < 1$ ($0 <$ mes M < 1)
ou bien : $\mathcal{P}_1 = 1$, $0 = \mathcal{P}_2$ (mes M $= 0$, p. ex. : M est vide).

ce qui montre bien que $x_\omega(.)$ n'est nullement définie par sa loi temporelle (le cas $\mathcal{P}_1 = 1$, $\mathcal{P}_2 = 0$ est particulièrement frappant).

2,23. Pour obvier aux conséquences de cet état de choses, on a proposé plusieurs remèdes qui, tous, reviennent à élargir au delà de $\mathcal{F}'_T$ le corps des sous-ensembles mesurables de $\mathcal{X}$. Je vais en citer quelques-uns, mais il est bien clair, après ce que je viens de dire, que ces remèdes ont nécessairement un certain arbitraire : si vraiment *la seule* donnée est la loi temporelle, les sous-ensembles de $\mathcal{X}$ qui n'appartiennent pas à $\mathcal{F}'_T$ peuvent, en général, être probabilisés de plusieurs façons différentes, toutes également compatibles avec la loi temporelle, et le choix entre les diverses définitions possibles ne saurait résulter que d'un coup de force impossible à justifier rationnellement.

2,231. Soit donc V un sous-ensemble de $\mathcal{X}$ qui n'appartienne pas à $\mathcal{F}'_T$; il y a des ensembles de $\mathcal{F}'_T$ qui contiennent V (par exemple, $\mathcal{X}$ lui-même); si la borne inférieure des mesures de ces ensembles (c'est-à-dire la P_3-mesure extérieure de V) vaut 1 :

$$\overline{P}_3(V) \equiv \inf_{M \in \mathcal{F}'_T,\ V \subset M} [\text{mes } M] = 1$$

on peut décider, en quelque sorte, de borner son attention aux éléments de V. Pour cela, on considère la famille $\mathcal{F}_V$ des sous-ensembles H de $\mathcal{X}$ qui sont de la forme :

$$H \equiv [F \cap V] \cup [G \cap (\mathcal{X} - V)]$$

F et G appartenant à $\mathcal{F}'_T$. On montre que cette famille est un corps borélien qui comprend $\mathcal{F}'_T$ et que, si l'on pose :

$$P_V(H) \equiv P_3(F)$$

la fonction d'ensemble P_V est une probabilité définie sur $\mathcal{F}_V$, et identique, sur $\mathcal{F}'_T$, à P_3. On a donc réussi à étendre la famille des ensembles probabilisés de $\mathcal{X}$ [en particulier, alors que $P_3(V)$ n'existe pas, on a $P_V(V) = 1$], et l'on peut fréquemment opérer cette extension d'une manière intuitivement satisfaisante. Une telle extension s'appelle « extension normale » (« standard extension » de Doob).

Ainsi, supposons que la loi temporelle ait la forme particulièrement simple :

$$P_3[\mathbf{x}(t_1) = 0, \ldots, \mathbf{x}(t_n) = 0] \equiv 1.$$

L'ensemble Γ formé de la seule et unique fonction $x(t) \equiv 0$ n'appartient pas à $\mathcal{F}'_T$ mais sa mesure extérieure vaut 1; on peut donc appliquer la théorie ci-dessus, ce qui donne non seulement

$$P_\Gamma[\mathbf{x}(t_1) = 0, \ldots, \mathbf{x}(t_n) = 0] = 1 \tag{5a}$$

mais encore

$$P_\Gamma[\mathbf{x}(t) = 0 \text{ pour tout } t] = 1 \tag{5b}$$

ce qui est intuitivement satisfaisant. Mais l'arbitraire d'une telle extension normale apparaît immédiatement si l'on remarque que le complémentaire de Γ, $\mathcal{X} - \Gamma$, c'est-à-dire l'ensemble $\mathcal{X}$ privé de la seule fonction identiquement nulle, a, lui aussi, une mesure extérieure égale à 1, ce qui conduit à écrire

$$P_{\mathcal{X}-\Gamma}[\mathbf{x}(t_1) = 0, \ldots, \mathbf{x}(t_n) = 0] = 1 \tag{6a}$$

$$P_{\mathcal{X}-\Gamma}[\mathbf{x}(t) = 0 \text{ pour tout } t] = 0 \tag{6b}$$

Le fait que (5) soit intuitivement agréable, et (6) paradoxal n'est évidemment pas, en soi, une raison suffisante pour choisir entre eux.

2,232. Une autre méthode, proposée par Kakutani, consiste à utiliser les propriétés de la topologie que l'on induit sur $\mathcal{X}$ en y définissant comme ouverts les ensembles de la forme :

$$a_1 < x(t_1) < b_1, \ldots, a_n < x(t_n) < b_n.$$

Elle aboutit à définir une famille $\mathcal{F}_K$ d'ensembles mesurables qui, sans aucun doute, contient tout ensemble ayant un intérêt pratique; mais il n'est nullement certain que cette méthode soit

la seule possible ni, par conséquent, que la probabilité des événements de $\mathcal{F}_K$ soit univoquement définie.

2,233. On a proposé aussi d'abandonner la loi temporelle comme point de départ de la théorie et de la remplacer par des données du type suivant :

$$\Phi_1 [a_i, b_i; \alpha_i, \beta_i; i = 1, ..., n] = \Pr [\Omega_{i=1}^{n} M_i]$$

avec

$M_i \equiv \{x(.) : a_i < x(t) < b_i$ pour tout t de l'intervalle $(\alpha_i, \beta_i)\}$.

On peut, à partir de là, édifier une théorie exactement parallèle à celle de la loi temporelle : on considère la famille $\mathcal{B}_1$ des ensembles M_i, sur laquelle est définie la fonction d'ensemble Φ_1, puis le corps borélien minimal $\mathcal{B}_2$ engendré par $\mathcal{B}_1$ et la probabilité Φ_2 univoquement définie sur $\mathcal{B}_2$ comme prolongement de Φ_1, enfin le corps borélien $\mathcal{B}'_2$ et la probabilité Φ'_2 résultant de la complétion de $\mathcal{B}_2$ et Φ_2. A la différence de $\mathcal{F}'_T$, $\mathcal{B}'_2$ contient tout événement pratiquement intéressant [p. ex., $\inf_{\mathcal{T}} \mathbf{x}(t)$ et $\sup_{\mathcal{T}} \mathbf{x}(t)$ sont *toujours* des variables aléatoires parfaitement définies sur $\mathcal{B}'_2$]. Mais cette théorie, si séduisante soit-elle, n'est au fond qu'un aveu d'impuissance : elle ne résout les problèmes de la théorie des fonctions aléatoires qu'en exigeant des données beaucoup plus abondantes que la seule loi temporelle; en outre, alors que, pratiquement, la loi temporelle est en général assez facile à obtenir, la fonction Φ_1 n'est connue que dans des cas simples, et seulement comme aboutissement d'une longue chaîne de déductions.

2,234. On peut donc, me semble-t-il, conclure ainsi le début de cette enquête : la loi temporelle d'une fonction aléatoire est la propriété qui s'offre le plus immédiatement à l'étude, mais sa donnée est insuffisante pour l'étude d'un grand nombre de propriétés intéressantes; une fonction aléatoire n'est donc que très incomplètement définie par la loi temporelle.

2,24. Il y a pourtant une propriété très importante qui ne dépend que de la loi temporelle : en effet, si l'on a, pour tout n et quels que soient les t_i $(t_1 < t_2 < ... < t_n < t_{n+1})$:

$$\begin{aligned} \mathrm{P}_3 [\mathbf{x}(t_{n+1}) \leqslant \lambda \,|\, \mathbf{x}(t_1) = a_1, ..., \mathbf{x}(t_n) = a_n] \\ = \mathrm{P}_3 [\mathbf{x}(t_{n+1}) \leqslant \lambda \,|\, \mathbf{x}(t_n) = a_n] \end{aligned} \qquad (7)$$

on dit que $\mathbf{x}(.)$ est « markoffienne » ou « du type de Markov »; il faut remarquer que, bien souvent, on a $P_3[\mathbf{x}(t_i) = a_i] = 0$, de sorte que les probabilités conditionnelles de (7) ont une définition quelque peu symbolique; à cette restriction près, on peut dire que, pour une fonction markoffienne, la donnée du présent rend l'avenir et le passé mutuellement indépendants. Ceci explique le rôle important que jouent, dans la théorie des fonctions de ce type, les « probabilités de transition » :

$$T(s, a; t, A) \equiv P_3[\mathbf{x}(t) \in A \mid \mathbf{x}(s) = a] \quad (s < t)$$

(où A est un intervalle semi-ouvert à gauche); on a notamment la relation dite « de Chapman-Kolmogorov », qui traduit le théorème des probabilités composées :

$$T(s, a; t, A) = \int_{-\infty}^{\infty} T(u, b; t, A)\ T(s, a; u, db) \quad (s < u < t).$$

L'une des applications les plus importantes de la théorie des fonctions markoffiennes se trouve dans la théorie statistique de la diffusion : on considère un volume V, $\mathbf{x}(t)$ est le nombre de particules intérieures à V à l'instant t, et, moyennant des hypothèses convenables, $\mathbf{x}(t + \delta)$ ne dépend pas, si $x(t)$ est connu, des valeurs prises par $\mathbf{x}(.)$ avant l'instant t.

Il faut soigneusement se garder de croire qu'une fonction aléatoire du type de Markov garde cette propriété quand on groupe d'une manière quelconque les états qu'elle peut prendre; ainsi, soit $\mathbf{x}(.)$ une fonction markoffienne, et posons :

$$y(t) = \begin{cases} 1 \text{ si } x(t) > a \\ 0 \text{ si } x(t) \leqslant a. \end{cases}$$

La connaissance de la loi temporelle de $\mathbf{y}(.)$ résoudrait fort simplement des problèmes très importants relatifs aux « temps de récurrence » et « temps de passage » de $\mathbf{x}(.)$. Mais il est essentiel de bien voir que, sauf exception, $\mathbf{y}(.)$ n'est *pas* une fonction markoffienne; plus généralement, on n'a pas, pour $t_1 < t_2 < t_3$:

$$P_3[\mathbf{x}(t_3) \in A_3 \mid \mathbf{x}(t_2) \in A_2, \mathbf{x}(t_1) \in A_1] = P_3[\mathbf{x}(t_3) \in A_3 \mid \mathbf{x}(t_2) \in A_2].$$

2,3. On voit donc que l'étude d'une fonction aléatoire à partir de sa loi temporelle, ou, comme on dit, « du point de vue de Slutsky », n'est pas tellement intéressante. Aussi n'y a-t-il, à ma

connaissance du moins, qu'un seul exemple important de fonction aléatoire ainsi étudiée, c'est la théorie du mouvement brownien.

2,31. Sans entrer dans trop de détails, je vais reprendre la définition et quelques propriétés de la fonction aléatoire du mouvement brownien linéaire (f. a. de Wiener-Lévy) telles qu'on les trouve, par exemple, dans le livre bien connu de P. Lévy. Des considérations heuristiques conduisent à admettre que, si $\mathbf{x}(.)$ est une des coordonnées cartésiennes d'une particule soumise au mouvement brownien, l'accroissement $\mathbf{x}(t_2) - \mathbf{x}(t_1)$ de $\mathbf{x}(.)$ dans un intervalle (t_1, t_2) quelconque est une variable gaussienne de moyenne nulle et écart-type $\sqrt{t_2 - t_1}$ indépendante des valeurs prises par $\mathbf{x}(.)$ aux instants antérieurs à t_1 et même de $\mathbf{x}(t_1)$. Si on pose en outre $\mathbf{x}(0) \equiv 0$, la loi temporelle de $\mathbf{x}(.)$ est ainsi entièrement définie. Pour le montrer, bornons-nous aux valeurs > 0 de t, et considérons une suite quelconque de valeurs de t :

$$t_0 = 0, t_1, t_2, \ldots, t_n, \ldots$$

Si on a déterminé les valeurs $x(t_1), \ldots, x(t_{p-1})$ on peut, pour la détermination de $x(t_p)$, se trouver devant deux situations quelque peu différentes :

ou bien il existe parmi les t_i avec $i < p$ une valeur au moins qui est $> t_p$; alors, t_p est encadré par deux valeurs t_p' et t_p'' :

$$t_p' \equiv t_j,\ j < p;\ t_p'' \equiv t_k,\ k < p;\ t_p' < t_p < t_p'';$$

$x(t_p')$ et $x(t_p'')$ sont donc connus, les variables aléatoires $\mathbf{x}(t_p'') - \mathbf{x}(t_p)$ et $\mathbf{x}(t_p) - \mathbf{x}(t_p')$ sont mutuellement indépendantes, gaussiennes, de moyennes nulles et d'écarts-types respectifs

$$\sqrt{t_p'' - t_p} \text{ et } \sqrt{t_p - t_p'};$$

il en résulte immédiatement que la distribution *conditionnelle* de $\mathbf{x}(t_p)$, si $x(t_1), \ldots, x(t_{p-1})$ sont données, est celle d'une variable aléatoire gaussienne de moyenne

$$\frac{x(t_p'')\,(t_p - t_p') + x(t_p')\,(t_p'' - t_p)}{t_p'' - t_p'}$$

et écart-type :

$$\sqrt{(t_p - t_p')\,(t_p'' - t_p)\,/(t_p'' - t_p')};$$

ou bien il n'existe pas de $t_i > t_p$ avec $i < p$; alors, si t_p' est la plus grande des abscisses t_i avec $i < p$, la distribution *conditionnelle*

de $\mathbf{x}(t_p)$ si $x(t_1), \ldots, x(t_{p-1})$ sont données est celle d'une variable aléatoire gaussienne de moyenne $x(t_p')$ et d'écart-type $\sqrt{t_p - t_p'}$.

Il suffit de procéder ainsi de proche en proche pour déterminer la distribution conjointe des $\mathbf{x}(t_i)$, c'est-à-dire puisque les t_i sont quelconques, la loi temporelle de $\mathbf{x}(.)$ [accessoirement, on voit immédiatement que $\mathbf{x}(.)$ est du type de Markov]. Tout ceci s'applique évidemment au cas d'une suite t_i partout dense, par exemple la suite des nombres rationnels. Lévy démontre alors un théorème d'une grande importance : sur une telle suite dense, fixée une fois pour toutes, la fonction aléatoire $\mathbf{x}(.)$ est presque sûrement continue :

$$P_3\Big[\lim_{|t_j - t_i| \to 0} \mathbf{x}(t_j) = \mathbf{x}(t_i)\Big] = 1.$$

Ceci l'autorise alors à *compléter* la définition de $\mathbf{x}(.)$ par le *postulat* (qui, dans le cas actuel, est empiriquement justifié) que $\mathbf{x}(.)$ est presque sûrement continue partout, c'est-à-dire que, pour une valeur t n'appartenant pas à la suite t_i, il *définit* $\mathbf{x}(t)$ par la relation

$$x(t) = \lim_{t_i \to t} x(t_i);$$

car le théorème ci-dessus garantit que, sauf un événement de probabilité nulle, cette définition n'entraîne aucune contradiction. Au fond, le théorème de Lévy montre que, dans l'ensemble $\mathcal{X}$ de toutes les fonctions de t, le sous-ensemble des fonctions continues, qui n'est pas P_3-mesurable, a cependant une P_3-mesure extérieure égale à 1, de sorte qu'il existe une (et, essentiellement, une seule) extension normale partout continue. L'arbitraire qui s'attache, en principe, aux extensions normales est ici évité par un appel à l'empirisme : la fonction cherchée représente un phénomène physique qui est continu, la continuité de cette fonction est donc une *donnée supplémentaire* qui, jointe à la loi temporelle, détermine la fonction. C'est ainsi, par exemple, que l'on connaît, pour la fonction du mouvement brownien linéaire, la distribution du maximum et du minimum sur un intervalle donné, ou de la distance de deux extréma successifs. Ces résultats ont d'ailleurs été utilisés dans l'étude du critère d'ajustement de Kolmogorov-Smirnov.

3. Second point de vue : définition directe.

3,0. Si, comme nous venons de le voir, l'étude d'une fonction

aléatoire à partir de sa loi temporelle seule s'avère nécessairement incomplète, il faut en chercher la cause dans le fait que, par cette méthode, on part d'un ensemble fonctionnel extrêmement vaste, trop vaste certes pour les applications pratiques, ensemble dans lequel l'additivité dénombrable postulée n'est à peu près d'aucun secours. Aussi, en pratique, est-ce le plus souvent d'une tout autre façon que sont définies les fonctions aléatoires que l'on utilise.

3,1. Rappelons d'abord la définition d'une *variable* aléatoire. On dit communément qu'une variable aléatoire est « une grandeur dont les valeurs sont déterminées par le hasard »; grandeur signifie nombre, et on sait parfaitement que, si l'on parle de quelque chose qui est *variable*, surtout si ses valeurs sont déterminées de telle ou de telle façon, ce quelque chose est une *fonction*; puisqu'on parle de hasard, il s'agit d'une fonction définie sur un espace probabilisé, d'où la définition connue : une variable aléatoire est une fonction à valeurs réelles, définie sur un espace probabilisé, et mesurable; ou, plus explicitement, (E, $\mathcal{F}$, P) étant les éléments constitutifs d'un espace probabilisé, la variable aléatoire **v** est définie par une correspondance entre les éléments de E et les nombres réels :

$$e \rightarrow v\,(e)$$

telle que chacun des sous-ensembles de E : $\{e : v(e) \leqslant a\}$, a étant un nombre réel quelconque, appartienne à $\mathcal{F}$, de sorte que $\mathrm{P}\,[\mathbf{v}(e) \leqslant a]$ ait toujours une valeur bien définie.

3,2. On ne procède pas autrement pour définir une fonction aléatoire; la seule différence est que l'on applique E non plus sur l'ensemble des réels, mais sur un ensemble de fonctions.

Historiquement, Fréchet a été l'un des premiers à mettre ce point de vue en œuvre d'une manière explicite. Il considérait par exemple comme ensemble probabilisé une certaine population, au sens statistique, d'individus humains, E, à chaque élément e de laquelle il attachait une fonction (périodique de période 2π), $\varphi_e(.)$, à savoir le 2$^{\mathrm{d}}$ membre de l'équation polaire, $r = \varphi_e(\omega)$, du contour crânien de l'individu. Il définissait bien ainsi une fonction aléatoire $\boldsymbol{\varphi}(.)$ dont il pouvait étudier diverses propriétés; par exemple, il pouvait déterminer l'analogue de la loi temporelle

$$\mathrm{Pr}\,[\boldsymbol{\varphi}(t_i) \leqslant a_1, \ldots, \boldsymbol{\varphi}(t_n) \leqslant a_n];$$

ou encore, déterminer les coefficients de Fourier, qui sont évidemment des variables aléatoires :

$$\varphi(t) \equiv \mathbf{a}_0 + \Sigma\, \mathbf{a}_i \sin (2\pi/m)\, t,$$

et chercher, par exemple, les moyennes et les écarts-type de ces coefficients. Au moment où Fréchet commençait ces recherches, un point de vue entièrement semblable était depuis longtemps implicite en mécanique statistique; ici, l'espace probabilisé E n'est autre que l'espace des phases muni d'une mesure $\mu(e)$, dont le choix d'ailleurs ne fait d'ordinaire pas l'objet d'une discussion fort approfondie. A chaque point e de cet espace correspond une fonction (vectorielle) du temps, $\vec{x}_e(.)$, qui décrit l'évolution temporelle du système observé, sous l'hypothèse que, à l'origine des temps ($t = 0$) ce système est représenté par le point e; la correspondance

$$e \to \vec{x}_e\,(.)$$

définit la fonction aléatoire $\vec{\mathbf{x}}(.)$. La physique expérimentale conduit en fait à observer des variables aléatoires du type :

$$e \to \overline{\mathrm{A}}_{e,t} \equiv 1/\mathrm{T} \int_t^{t+\mathrm{T}} \mathrm{A}\,[\vec{x}_e(s)]\, ds$$

et l'un des principaux parmi les problèmes généraux de la mécanique statistique est précisément l'étude des relations entre les grandeurs de ce type, les seules directement observables, et celles du type

$$\mathcal{E}\,\Phi_t \equiv \int_{\mathrm{E}} \mathrm{A}\,[\vec{\mathbf{x}}_e\,(t)]\, d\mu_e,$$

que l'on suppose indépendantes de t, et qui sont en fait les seules directement accessibles au calcul.

3,31. D'une manière générale, la définition d'une fonction aléatoire suppose donnés :

1°) un espace probabilisé (catégorie d'épreuves) (E, $\mathcal{F}$, P);

2°) un ensemble Λ de fonctions $\lambda(.)$; par exemple, Λ est l'ensemble des fonctions continues d'une variable réelle, ou des fonctions à valeurs entières, etc;

3°) une correspondance qui, à chaque élément e de E attache un élément λ_e de E : $e \to \lambda_e(.)$. (7)

[Chacun de ces éléments λ_e est appelé une « réalisation » de la fonction aléatoire $\mathbf{x}(.)$].

On peut alors considérer des sous-ensembles probabilisés de Λ, à savoir, les images par (7) des sous-ensembles de E qui appartiennent à $\mathcal{F}$, les images ayant, par définition, les mêmes probabilités que les originaux.

3,32. En principe, le procédé ainsi décrit est encore plus général que celui qui part de la loi temporelle, car celui-ci s'obtient simplement en prenant pour E, $\mathcal{F}$ et P respectivement $\mathcal{X}$, $\mathcal{F}'_T$ et P_3, et pour la correspondance $e \to \lambda_e(.)$ l'identité. Il semble donc, à première vue, que la formalisation actuelle soit une machine infernale aussi inutile qu'encombrante.

En pratique toutefois, il s'introduit très souvent une énorme simplification : on peut prendre pour E un ensemble contenant seulement une infinité dénombrable d'éléments. La propriété d'addivité complète donne alors un contrôle total des événements probabilisés, et permet, du moins en principe, de résoudre les problèmes les plus compliqués. En particulier, la loi temporelle est déduite de la définition de la fonction, au lieu de servir de base à celle-ci, et le corps $\mathcal{F}'_T$ décrit plus haut ne joue plus, en général, aucun rôle spécial.

3,33. L'espace fonctionnel Λ est d'ordinaire un espace distancié : $\lambda(.)$ et $\mu(.)$ étant 2 fonctions quelconques de Λ, il leur correspond une distance $d(\lambda, \mu)$ qui permet de définir une convergence dans Λ.

La notion de « convergence presque sûre » est alors très importante; par exemple, si l'on considère une famille de fonctions aléatoires dépendant d'un paramètre réel ξ, c'est-à-dire une correspondance du type :

$$(e, \xi) \to x_e(.\,;\xi) \quad [e \in E,\ x_e(.\,;\xi) \in \Lambda],$$

on dit que cette famille converge presque sûrement vers la fonction aléatoire $\mathbf{y}(.)$ lorsque $\xi \to \xi_0$ s'il existe une fonction $\varphi_e(\eta) > 0$ et définie pour tout $\eta > 0$, telle que le sous-ensemble de E :

$$\{e : \inf_{|\xi - \xi_0| < \varphi_e(\eta)} d[x_e(.\,;\xi), y_e(.)] < \eta\}$$

est mesurable et de probabilité 1 quel que soit η; autrement dit :

$$P\{\lim_{\xi \to \xi_0} d[\mathbf{x}(.\,;\xi), \mathbf{y}(.)] = 0\} = 1.$$

Deux cas particuliers importants sont les suivants :

a) t étant donné, on prend $d(\lambda, \mu) \equiv |\lambda(t) - \mu(t)|$; il y correspond

la notion de « convergence locale presque sûre » :

$$P[\lim_{\xi \to \xi_0} \mathbf{x}(t); \xi) = \mathbf{y}(t)] = 1;$$

b) l'intervalle $\mathcal{J}$ étant donné, on prend $d(\lambda, \mu) = \sup_{\mathcal{J}} |\lambda(t) - \mu(t)|$;

il y correspond la notion de « convergence uniforme presque sûre » :

$$P[\lim_{\xi \to \xi_0} \sup_{t \in \mathcal{J}} |\mathbf{x}(t; \xi) - \mathbf{y}(t)| = 0] = 1.$$

Les définitions de convergence entraînent évidemment des définitions analogues pour la continuité; ainsi, on dit que la fonction aléatoire $\mathbf{x}(.)$ est presque sûrement localement continue au point t si

$$P[\lim_{\xi \to 0} \mathbf{x}(t + \xi) = \mathbf{x}(t)] = 1$$

et qu'elle est presque sûrement (uniformément) continue dans l'intervalle $\mathcal{J}$ si

$$P[\lim_{\xi \to 0} \sup_{t \in \mathcal{J}} |\mathbf{x}(t + \xi) - \mathbf{x}(t)| = 0] = 1.$$

Nous verrons tantôt que la continuité locale presque sûre en tout point d'un intervalle est une propriété beaucoup plus faible que la continuité presque sûre dans cet intervalle. Il y a ici une différence très nette entre le comportement des fonctions certaines et celui des fonctions aléatoires.

4. Exemples.

4,1. Un premier exemple, bien connu mais particulièrement utile, est fourni par les fonctions aléatoires liées à la loi de Poisson. Il s'agit de fonctions à valeurs entières $\geqslant 0$ d'une variable réelle. On arrive à leur définition précise à partir des hypothèses que voici :

a) on admet que des « survenances » sont susceptibles de se produire à des instants quelconques, de sorte que, l'intervalle $\mathcal{J}$ étant donné, le nombre des survenances de $\mathcal{J}$ est une variable aléatoire à valeurs entières $\geqslant 0$, $\mathbf{y}(\mathcal{J})$;

b) si $\mathcal{J}$ et $\mathcal{G}$ sont disjoints, $\mathbf{y}(\mathcal{J})$ et $\mathbf{y}(\mathcal{G})$ sont mutuellement indépendantes;

c) si $\mathcal{J}$ et $\mathcal{G}$ ont la même longueur, $\mathbf{y}(\mathcal{J})$ et $\mathbf{y}(\mathcal{G})$ ont la même distribution ;

d) si $|\mathcal{J}|$ est la longueur de $\mathcal{J}$ on a :

$$\Pr[\mathbf{y}(\mathcal{J}) = 1] = \lambda|\mathcal{J}| + 0(|\mathcal{J}|);\ \Pr[\mathbf{y}(\mathcal{J}) > 1] = 0(|\mathcal{J}|).$$

Un raisonnement classique conduit alors à voir que $\mathbf{y}(\mathcal{J})$ est une variable de Poisson de moyenne $\lambda|\mathcal{J}|$:

$$\Pr[\mathbf{y}(\mathcal{J}) = n] = \exp(-\lambda|\mathcal{J}|)\,(\lambda|\mathcal{J}|)^n/n! \qquad (8)$$

Soit alors t_0 un instant fixe, δ la longueur du plus grand des intervalles (t_0, t_1) qui ne contiennent aucune survenance :

$$\Pr[\boldsymbol{\delta} > a] \equiv \Pr[\mathbf{y}(t_0, t_0 + a) = 0] = \exp(-\lambda a),$$

et ce raisonnement subsiste, inchangé, si l'on cherche la distribution de $\boldsymbol{\delta}$ conditionnelle au fait que t_0 est précisément l'abscisse d'une survenance. Et l'on arrive ainsi à la *définition* constructive et explicite de la fonction aléatoire $\mathbf{x}(.)$ liée à la loi de Poisson d'intensité λ :

espace E : l'espace à une infinité dénombrable de coordonnées réelles $\geqslant 0$:

$$\ldots, a_{-n}, \ldots, a_{-1}, a_1, \ldots, a_n, \ldots;$$

famille $\mathcal{F}$: le corps borélien engendré par les ensembles de la forme $a_i \leqslant \alpha_i$;

probabilité P : définie sans aucune ambiguïté par

$$\Pr[\mathbf{a}_{i_1} > u_{i_1}, \ldots, \mathbf{a}_{i_p} > u_{i_p}] = \exp[-\lambda \Sigma_1^p u_{i_k}];$$

ensemble fonctionnel Λ : l'ensemble de toutes les fonctions d'une variable réelle à valeurs entières $\geqslant 0$;

fonction aléatoire $\mathbf{x}(.)$ définie ainsi : au point $e \equiv (\ldots, a_{-n}, \ldots, a_n, \ldots)$ de E correspond la fonction $x_e(.)$ de Λ telle que, si l'on pose $A_r = a_1 + \ldots + a_r$, $A_{-r} = a_{-1} + \ldots + a_{-r}$ $(r > 0)$, on ait :

$$x_e(t) = 0 \quad \text{si} \quad -A_{-1} < t \leqslant A_1;$$

$$x_e(t) = -1 \quad \text{si} \quad -A_{-2} < t \leqslant -A_{-1};\quad x_e(t) = 1 \quad \text{si} \quad A_1 < t \leqslant A_2;$$

$$x_e(t) = -p \quad \text{si} \quad -A_{-p-1} < t \leqslant -A_p;\ x_e(t) = p \quad \text{si} \quad A_p < t \leqslant A_{p+1};$$

...A partir de là, la loi temporelle de $\mathbf{x}\ldots(.)$ s'écrit sans difficulté.

Il est intéressant de montrer que cette fonction aléatoire est partout presque sûrement localement continue; t_0 étant fixé, on a en effet, quel que soit $\varepsilon > 0$

$$\Pr[\lim_{t \to t_0} \mathbf{x}(t) = \mathbf{x}(t_0)] \geqslant \Pr[\mathbf{x}(t+\varepsilon) - \mathbf{x}(t-\varepsilon) = 0]$$

$$= \exp(-2\lambda\varepsilon) \geqslant 1 - 2\lambda\varepsilon,$$

donc

$$\Pr[\lim_{t \to t_0} \mathbf{x}(t) = \mathbf{x}(t_0)] = 1.$$

Mais $\mathbf{x}(.)$ n'est cependant presque sûrement continue sur aucun intervalle, car :

$$\Pr[\mathbf{x}(.) \text{ continue sur } (t_0, t_1)] = \Pr[\mathbf{x}(t_1) - \mathbf{x}(t_0) = 0]$$

$$= \exp[-\lambda(t_1 - t_0) < 1.$$

On a ainsi un exemple très simple de la différence entre ces deux notions de continuité stochastique.

4,2. Une classe très intéressante de fonctions aléatoires est formée par les fonctionnelles d'une fonction liée à la loi de Poisson.

4,21. Ainsi, l'étude de l'effet de grenaille pur dans l'émission thermo-ionique conduit à admettre que les arrivées d'électrons sur l'anode sont des survenances poissonniennes, la réponse fournie au temps $t + \theta$ par l'appareil enregistreur à une survenance au temps t étant $R(\theta)$ et les réponses à diverses survenances étant purement et simplement additives; on doit alors considérer la fonction aléatoire $\mathbf{r}(.)$ définie, sur le même espace probabilisé que ci-dessus, par :

$$\mathbf{r}(t) = \sum_{\mathbf{A}_i \leqslant t} R(t - \mathbf{A}_i)$$

On doit évidemment démontrer que cette série à termes aléatoires est presque sûrement convergente, ce qui est facile si $R(\theta)$ décroît suffisamment vite; la loi temporelle de $\mathbf{r}(.)$ est alors facile à calculer.

Remarquons que $\mathbf{r}(.)$ peut fort bien n'être ni à valeurs entières ni presque sûrement continue : sa définition à partir de la loi temporelle conduirait à d'énormes difficultés.

4,22. Deux problèmes très différents, l'étude de l'effet de scintillation des cathodes émissives en atmosphère gazeuse et celle de la régularité des fils formés de fibres conduisent à des fonctionnelles plus compliquées de la fonction $\mathbf{x}(.)$.

Ici, on prend pour E un espace à une double infinité de coordonnées réelles $\geqslant 0$:

$$\ldots, a_{-n}, \ldots, a_{-1}, a_1, \ldots, a_n, \ldots$$

$$\ldots, l_{-n}, \ldots, l_{-1}, l_1, \ldots, l_n, \ldots$$

les l_i étant d'ailleurs toutes > 0; le corps $\mathcal{F}$ est, cette fois, engendré par les ensembles de la forme $a_i \leqslant u_i$ ou $l_j \leqslant v_j$ et la fonction P est définie par :

$$P\,[\mathbf{a}_{i_1} \leqslant u_{i_1}, \ldots, \mathbf{a}_{i_p} \leqslant u_{i_p}, \mathbf{l}_{k_1} \leqslant v_{k_1}, \ldots, \mathbf{l}_{k_n} \leqslant v_{k_n}] =$$

$$\Pi_{r=1}^{p}\,[1 - \exp(-\lambda u_{i_r})]\;\Pi_{s=1}^{n} F(v_{k_s}),$$

F étant une certaine fonction de répartition satisfaisant à $F(+o) = 0$.

Ici, on considère qu'à la survenance d'abscisse A_i est attaché un « intervalle d'adsorption » ou bien une « fibre » de longueur l_i, et on s'intéresse au nombre de molécules adsorbées, ou de fibres, présentes à l'instant t, c'est-à-dire que, si l'on pose

$$n_i(t) = \begin{cases} 1 \text{ si } A_i < t \leqslant A_i + l_i \\ 0 \text{ autrement} \end{cases}$$

la fonction étudiée, $\mathbf{n}(.)$, est définie par :

$$n(t) = \Sigma_{-\infty}^{\infty} n_i(t).$$

Ici encore, $\mathbf{n}(.)$ est la somme d'une série à termes aléatoires. On en profite pour chercher, par exemple, la distribution de la variable aléatoire [définie sur $(E, \mathcal{F}, P)$] :

$$1/T \int_t^{t+T} \mathbf{n}(s)\, ds$$

dont l'existence ne fait ici l'objet d'aucun problème, alors qu'elle ne serait nullement garantie à partir de la seule donnée de la loi temporelle.

4,3. Un exemple un peu différent est fourni par les processus de renouvellement et ceux qui s'y rattachent.

4,31. Imaginons que, dans un établissement industriel, on utilise un type d'appareil susceptible de « vivre » un certain temps, d'ailleurs aléatoire. Il existe quantité de tels appareils, depuis le pneu d'auto jusqu'au tube électronique. On sait qu'il y a constamment

en service N tels appareils d'un type donné, que la durée de vie d'un appareil est une aléatoire dont la fonction de répartition est $F(t)$, que tout appareil qui meurt est immédiatement remplacé par un autre, d'âge 0, et enfin que, tant dans une même lignée que dans des lignées parallèles, les durées de vie des diverses appareils sont mutuellement indépendantes.

On s'intéresse essentiellement à la fonction aléatoire

$$\mathbf{r}(.) = \mathbf{N}_x(.)/\mathrm{N},$$

$\mathbf{N}_x(t)$ étant, parmi les divers appareils qui existent au temps t, le nombre de ceux dont l'âge est $\leqslant x$. Quelques artifices spéciaux, notamment l'emploi de la 2e caractéristique de la fonction de répartition $\mathbf{r}.(t)$, permettent de ramener le problème au cas essentiel où $\mathrm{N} = 1$, puis à l'étude de la fonction aléatoire fondamentale $\boldsymbol{\alpha}(.)$, $\boldsymbol{\alpha}(t)$ étant l'âge de l'unique appareil en service au temps t [de sorte que $r_x(t) = 0$ is $x < \alpha(t)$ et $= 1$ si $x > \alpha(t)$]. Supposons, pour simplifier, que tous les appareils sont mis en service au temps 0 avec l'âge 0. On prend alors (pour $\mathrm{N} = 1$) comme espace probabilisé E l'espace à une infinité de coordonnées réelles $\geqslant 0$ $a_1, a_2, \ldots, a_n, \ldots$, le corps borélien $\mathcal{F}$ étant celui qu'engendrent les ensembles de la forme $a_i \leqslant u_i$ et la mesure P étant définie par :

$$\mathrm{P}\,[\mathbf{a}_{i_1} \leqslant u_{i_1}, \ldots, \mathbf{a}_{i_n} \leqslant u_{i_n}] = \Pi_{k=1}^{p}\,\mathrm{F}\,(u_{i_k}).$$

Si on pose encore $\mathrm{A}_n = a_1 + \ldots + a_n$, la fonction aléatoire $\boldsymbol{\alpha}(.)$ est définie ainsi : au point $e \equiv (a_1, \ldots, a_n, \ldots)$ de E correspond, dans l'ensemble des fonctions d'une variable réelle, la fonction

$$\alpha_e(t) = \begin{cases} t & \text{si } 0 \leqslant t < \mathrm{A}_1 \\ t - \mathrm{A}_1 & \text{si } \mathrm{A}_1 \leqslant t < \mathrm{A}_2 \\ t - \mathrm{A}_n & \text{si } \mathrm{A}_n \leqslant t < \mathrm{A}_{n+1} \end{cases}$$

La loi temporelle se déduit de la relation :

$$\mathrm{P}\,[\boldsymbol{\alpha}(t) \leqslant u] = \Sigma_{n=0}^{\infty}\,\mathrm{P}\,[\mathbf{A}_n < t, \mathbf{A}_{n+1} \geqslant t].\ \mathrm{P}\,[\mathrm{A}\,t - \mathbf{A}_n \leqslant u].$$

Ici encore, tout événement pratiquement intéressant appartient à $\mathcal{F}$.

4,32. Un processus analogue au précédent est utilisé dans l'étude de la croissance des populations. Ici, chaque individu peut, durant sa vie, engendrer un certain nombre de descendants qui, à leur tour, peuvent en engendrer d'autres et ainsi de suite. L'espace E

se décrit le plus aisément de proche en proche : l'individu initial, I_0, naît en $t = 0$; sa vie a une durée aléatoire $\mathbf{v}_0$ de fonction de répartition $F(\xi)$; l'instant $\mathbf{n}_{0,1}$ où il engendre son 1er descendant a lui-même une fonction de répartition $N_1(\xi)$ mais n'est pris en considération que s'il survient durant la vie de l'individu; de même l'instant $\mathbf{n}_{0,2}$ où il engendre son 2e descendant à une fonction de répartition *conditionnelle* $N_2(\xi|n_{0,1})$ et n'est, lui non plus, pris en considération que si $n_{0,2} < v_0$; et ainsi de suite. On procède de même pour chaque descendant. La situation devient vite assez compliquée, mais l'essentiel est que l'espace E est, ici encore, adéquatement décrit par une infinité dénombrable de coordonnées réelles, ce qui garantit la mesurabilité des événements que l'on est amené à considérer.

4,4. On voit, sur ces quelques exemples, comment une définition directe des fonctions aléatoires permet de résoudre les problèmes auxquels se heurte la définition à partir de la loi temporelle. La théorie des fonctions aléatoires est de plus en plus appliquée par les praticiens : par exemple, en physique, à l'étude des rayons cosmiques ou à celle de l'effet du bruit de fond dans les télécommunications; en biologie, à l'étude de l'hérédité ou à celle des épidémies. On peut dire que, presque toujours, c'est la méthode directe qui est utilisée, mais d'une manière peu explicite.

Il semble donc que les praticiens choisissent d'instinct la méthode qui est, sinon la plus immédiate, du moins la plus puissante et la plus sûre. Mais il semble aussi qu'ils auraient avantage à prendre plus explicitement conscience de la manière dont la méthode directe résout des difficultés sérieuses au point d'avoir parfois troublé même les spécialistes de ces questions.

BIBLIOGRAPHIE SOMMAIRE

P. Lévy : *Processus stochastiques et mouvement brownien*; Paris (Gauthier-Villars) 1948.

J. L. Doob : *Stochastic processes*; New -York (Wiley) 1953.

R. Fortet et A. Blanc-Lapierre : *Fonctions aléatoires*; Paris (Masson) 1954.

Annales de la Société Scientifique de Bruxelles, Tome 70, nº II, p. 117-123

Intégration par analyse harmonique

NOTE DE

M. le Chanoine G. LEMAITRE

1. — Méthode d'intégration

Considérons une équation différentielle

$$L(y) = z$$

où $L(y)$ est une expression linéaire en y et ses dérivées par rapport à x et où z est une fonction déterminée de y et de x.

Nous étudierons en particulier le cas 1) où $L(y) = dy/dx$ et le cas 2) où $L(y) = d^2y/dx^2$.

Les fonctions $y(x)$ et $z(x)$ sont supposées être des fonctions périodiques de période 2π ayant une valeur moyenne nulle, représentables à une certaine approximation par des expressions de la forme

$$y(x) = \sum_{q=1}^{k} (a_q \cos qx + b_q \sin qx)$$

$$z(x) = \sum_{q=1}^{k} (A_q \cos qx + B_q \sin qx)$$

avec $b_k = B_k = 0$. Nous utiliserons successivement $k = 2, 4, 8, 16$, etc.

Dans ces conditions, l'équation différentielle établit une relation entre a_q, b_q et A_q, B_q. Dans le cas général cette relation s'écrit

$$a_q = L_q^1 A_q + L_q^2 B_q$$

$$b_q = L_q^3 A_q + L_q^4 B_q$$

dans le cas 1), elle se réduit à

$$a_q = -1/q\, B_q \quad b_q = 1/q\, A_q$$

et dans le cas 2) à

$$a_q = -1/q^2 A_q \quad B_q = -1/q^2 A_q$$

En écrivant $x_j = j\pi/k$; j prenant les $2k$ valeurs équidistantes $-k, -k+1, \ldots, k-2, k-1$ et $z_j = z(x_j)$, $y_j = y(x_j)$ on a

$$A_q = 1/k \sum_j z_j \cos qx_j, \quad A_k = 1/2k \sum_j z_j \cos kx_j$$

$$B_q = 1/k \sum_j z_j \sin qx_j$$

De ces relations on peut déduire une relation directe entre les y_j et les z_j en éliminant les a_q, b_q, A_q, B_q. On peut même trouver une relation donnant un $y(x)$ quelconque connaissant les $2k$ valeurs de $z(x + x_{j'})$ où les x_j, prennent $2k$ valeurs équidistantes.

On aura en général une relation de la forme

$$y(x) = \sum_{j'} L_{j'} z(x_j + x_{j'})$$

On trouve, en particulier dans le cas 1)

$$L_{j'} = S_{j'} = -1/k \sum_{q=1}^{k} 1/q \sin qx_j,$$

et pour le cas 2)

$$L_{j'} = C_{j'} = -1/k \sum_{q=1}^{k-1} 1/q^2 \cos qx_{j'} - \frac{1}{2k^3} \cos kx_{j'}$$

Si on détermine les j' en partant de zéro, on peut au moyen des z_j calculer les y_j.

Si on détermine les j' à partir de $j' = 1/2$, on peut au moyen des mêmes z_j calculer les $y_{j+1/2}$. On a ainsi le moyen de doubler la valeur de k utilisée en interpolant les valeurs intermédiaires.

Pour chaque valeur de k utilisée, on sera ramené à la résolution d'un système d'équations

$$y_j = \Sigma_{j'} L_{j'} z(x_j + x_{j'})$$

entre les $2k$ quantités y_j et les $2k$ quantités z_j (chacunes de somme nulle); les z_j étant des fonctions déterminées des y_j.

Lorsqu'il est aisé d'obtenir la dérivée de z par rapport à y, on peut améliorer une approximation obtenue en procédant par variation, on aura

$$\delta y_j = \sum_{j'} L_{j'} \delta z(x_j + x_{j'})$$

et

$$\delta z_j = (\partial z/\partial y_j) \delta y_j$$

Considérant $(\partial z / \partial y_j)$ comme des constantes, on sera ramené à un système d'équations linéaires en les δy_j.

Après avoir répété éventuellement ce procédé pour une même valeur de k, on calculera les y pour les valeurs intermédiaires et on aura une première approximation pour aborder le problème pour la valeur double de k.

Il est clair que la méthode peut être aisément étendue à la résolution d'un système d'équations différentielles.

Dans la pratique, il pourra y avoir lieu de tenir compte des propriétés de symétrie particulières au problème.

2. — Application numérique

En vue d'illustrer le fonctionnement de la méthode nous l'avons appliquée à la détermination d'une fonction périodique de période 2π, s'annulant pour $x = 0$ et satisfaisant l'équation différentielle

$$d^2y / dx^2 = -4 \sin y$$

dont la solution s'obtient aisément par les fonctions elliptiques.

Pour $k = 4$, on aura par raison de symétrie

$$z_0 = z_4 = 0$$

$$z_1 = -z_{-1} = -z_{-3} = z_3$$

$$z_2 = -z_{-2}$$

et des relations analogues pour les y.

On aura à résoudre les deux équations en y_1 et y_2

$$y_1 = -4 \left[(C_0 - C_4) \sin y_1 + (C_1 - C_3) \sin y_2 \right]$$

$$y_2 = -4 \left[2 (C_1 - C_3) \sin y_1 + (C_0 - C_4) \sin y_2 \right]$$

Les valeurs numériques sont données en appendice.

On obtient comme solution

$$y_1 = 126^{\circ}92$$

$$y_2 = 159^{\circ}58$$

On introduit ensuite des valeurs intermédiaires par

$$y_{1/2} = (C_{1/2} - C_{3/2} + C_{5/2} - C_{7/2})\ z_1 + (C_{3/2} - C_{5/2})\ z_2$$

$$y_{3/2} = (C_{1/2} + C_{3/2} - C_{5/2} - C_{7/2})\ z_1 + (C_{1/2} - C_{7/2})\ z_2$$

Si on avait jugé bon de faire le calcul pour $k=2$ on aurait trouvé

$$y_1 = 100^o$$
$$y_2 = 142^o$$

En passant au calcul pour $k = 8$ les quantités

$$y_{1/2},\ y_1,\ y_{3/2},\ y_2,$$

seront dorénavant désignées par

$$y_1,\ y_2,\ y_3,\ y_4.$$

Pour $k = 8$; les équations deviennent

$$y_i = -4 \sum_k c_{ik} \sin y_k$$

où les c_{ik} sont

$i \downarrow$ \ $k \rightarrow$				
	$C_0 - C_2 + C_6 - C_8$	$C_1 - C_3 + C_5 - C_7$	$C_2 - C_6$	$C_3 - C_5$
	$C_1 - C_3 + C_5 - C_7$	$C_0 - C_8$	$C_1 + C_3 - C_5 - C_7$	$C_2 - C_6$
	$C_2 - C_6$	$C_1 + C_3 - C_5 - C_7$	$C_0 + C_2 - C_6 - C_8$	$C_1 - C_7$
	$2\,(C_3 - C_5)$	$2\,(C_2 - C_6)$	$2\,(C_1 - C_7)$	$C_0 - C_8$

Portant dans le second membre les valeurs y_i obtenues on trouve $y_i + r_i$ et la résolution de l'équation aux variations conduit à adopter $y_i + \delta y_i$. On a trouvé

y_i	r_i	δy_i	$y_i + \delta y_i$
74°08	7°41	6°72	80°79
126°92	5°85	4°04	130°96
152°82	3°29	1°00	153°82
159°58	3°01	0°62	160°20

Les corrections δy_i s'obtiennent par la résolution d'un système linéaire de la forme

$$\sum_k a_{ik}\, \delta y_k = r_i$$

où

$$a_{ik} = \delta_{ik} - 4 \sum c_{ik} \cos y_k$$

Il n'a pas été jugé utile de faire une seconde approximation et on a passé directement au calcul des valeurs intermédiaires par les formules

$$y_{i-1/2} = -4 \sum_k c'_{ik} \sin y_k$$

avec

$c'_{ik} =$

$C_{1/2} - C_{3/2} + C_{13/2} - C_{15/2}$	$C_{3/2} - C_{5/2} + C_{11/2} - C_{13/2}$	$C_{5/2} - C_{7/2} + C_{9/2} - C_{11/2}$	$C_{7/2} - C_{9/2}$
$C_{1/2} - C_{5/2} + C_{11/2} - C_{15/2}$	$C_{1/2} - C_{7/2} + C_{9/2} - C_{15/2}$	$C_{3/2} + C_{7/2} - C_{9/2} - C_{13/2}$	$C_{5/2} - C_{11/2}$
$C_{3/2} - C_{7/2} + C_{9/2} - C_{13/2}$	$C_{1/2} + C_{7/2} - C_{9/2} - C_{15/2}$	$C_{1/2} + C_{5/2} - C_{11/2} - C_{15/2}$	$C_{3/2} - C_{13/2}$
$C_{5/2} + C_{7/2} - C_{9/2} - C_{11/2}$	$C_{3/2} + C_{5/2} - C_{11/2} - C_{13/2}$	$C_{1/2} + C_{3/2} - C_{13/2} - C_{15/2}$	$C_{1/2} - C_{15/2}$

On a obtenu

y_i	r_i	δy_i	y'_i	r'_i	$\delta y'_i$	y''_i	r''_i	Jacobi
42°75	0°72	0°49	43°24	— 0°18	— 0°06	43°18	0°09	43°19
80°79	0°45	—0°05	80°74	— 0°38	— 0°13	80°61	0°19	80°63
108°79	1°76	1°00	109°79	— 0°58	— 0°20	109°59	0°28	109°63
130°96	0°64	—0°36	130°60	— 0°77	— 0°27	130°34	0°37	130°41
143°00	2°94	1°70	144°70	— 0°95	— 0°34	144°36	0°45	144°46
153°82	1°17	—0°22	153°60	— 1°07	— 0°38	153°22	0°51	153°35
156°52	3°51	1°98	158°50	— 1°17	— 0°43	158°07	0°55	158°23
160°20	1°38	—0°16	160°04	— 1°18	— 0°42	159e62	0°56	159°75

Les calculs ont été faits au laboratoire de recherches numériques de l'université de Louvain par Mademoiselle A. Bartholomé.

La dernière colonne donne pour comparaison les valeurs obtenues par la série de Jacobi

$$y^{o} = 180/\pi \sum_{n=1}^{\infty} \frac{8q^{n+1/2}}{1+q^{2n+1}} \frac{\sin(2n+1)x}{2n+1}$$

pour $q = 0.205328$ vérifiant

$$1 + 2q + 2q^4 + 2q^9 = \sqrt{2}.$$

3. — Appendice, Valeurs numériques des coefficients

Pour les coefficients du 1) nous avons trouvé :

i	$k=4$	$k=8$		$k=16$
	$4S_{i/2}$	$8S_i$	$8S_{i+1/2}$	$16S_{2i}$
0	0	0	0	0
			— 1.333785	— 1.752739
1	— 1.294197	— 1.651493		— 1.223852
			— 1.252671	— 1.377153
2	— 1.442809	— 1.033706		— 1.103594
			— 1.126092	— 1.137896
3	— 0.899872	— 1.073169		— 0.935265
			— 0.830418	— 0.921471
4	— 0.666667	— 0.723810		— 0.754268
			— 0.743328	— 0.712787
5	— 0.692765	— 0.630360		— 0.568226
			— 0.433327	— 0.507536
6	— 0.442809	— 0.367039		— 0.379784
			— 0.352799	— 0.303984
7	— 0.087090	— 0.208684		— 0.190146
			— 0.039588	— 0.101246

pour le 2) nous avons :

i	$4C_{i/2}$	$8C_i$	$8C_{i+1/2}$	$16C_{2i}$
0	— 1.392361	— 1.519610		— 1.582392
			— 1.385168	— 1.353487
1	— 1.143176	— 1.081560		— 1.064038
			— 0.785038	— 0.807704
2	— 0.597289	— 0.559998		— 0.564665
			— 0.357165	— 0.344292
3	— 0.103253	— 0.144228		— 0.140976
			0.050803	0.041469
4	0.218750	0.207465		0.205858
			0.344009	0.349989
5	0.456807	0.474123		0.475663
			0.585569	0.581354
6	0.659789	0.669373		0.668397
			0.733410	0.735591
7	0.789622	0.782916		0.784039
			0.813580	0.812706
8	0.829861	0.823429		0.822588

Annales de la Société Scientifique de Bruxelles, Tome 70, n° II, p. 124-163

Calcul des probabilités dans certains espaces topologiques

PAR

René LAVENDHOMME

Professeur à l'Université Lovanium, Congo.

Chapitre II

MESURES FINIES

1. Définition des mesures finies

Nous supposerons comme au chapitre précédent que X est un espace topologique normal.

Les mesures que nous considérons sur X sont toujours définies à partir d'un espace vectoriel U de fonctions continues et bornées à valeurs dans un espace de Banach X'. Cet espace U satisfait à l'axiome suivant :

A) Si u est dans U, alors toute fonction continue v, telle que $\forall x$:

$$|v(x)| \leqslant |u(x)|,$$

est aussi dans U.

Dans ces conditions, on a vu au chapitre précédent que U détermine univoquement un espace de mesures M, le σ-anneau de parties mesurables pour toute mesure de M et celles de ces parties qui sont intégrables pour toute mesure de M.

U est muni d'une topologie pour laquelle il est limite inductive d'espaces de Banach. Lorsqu'on prend pour X' les réels, M n'est autre que le dual de cet U topologique. Un résultat important est que M est complet.

Le plus grand espace U que l'on puisse trouver est évidemment celui qui est formé de toutes les fonctions continues bornées : l'axiome A est trivialement vérifié.

Désignons par U_0 cet espace des fonctions continues bornées, et par M_0 l'espace des mesures correspondant.

Définition 1

Une mesure finie est une mesure qui appartient à l'espace M_0 défini à partir de l'espace U_0 de toutes les fonctions continues bornées.

C'est de l'espace des mesures finies que nous allons nous occuper dans ce chapitre. Nous allons d'abord énoncer quelques théorèmes qui résultent immédiatement des définitions.

Théorème 1. U_0 et M_0 sont des espaces de Banach.

En effet, U est toujours limite inductive d'espaces de Banach mais ici la famille des espaces de Banach U_{u_i}, (Chap. I, § 4), dont U est limite inductive, a un plus grand élément, l'espace U_0 lui-même, puisque U_0 est l'espace U_u avec $u = \text{const.}$ U_0 est donc un espace de Banach.

M_0 est aussi espace de Banach, puisqu'il est le dual de U_0 quand on prend pour X' l'espace des réels.

Théorème II. Tout ensemble ouvert est intégrable.

En effet, cela résulte du théorème VI du chapitre précédent puisque la caractéristique d'un ouvert est toujours inférieure à une fonction constante dans tout l'espace.

Remarque. En particulier l'espace X tout entier est intégrable, ce qui justifie l'appellation mesure finie adoptée pour les éléments de M_0.

Théorème III. Comme espaces topologiques, on a $U \subseteq U_0$ pour tout U vérifiant l'axiome A.

En effet, $U \subseteq U_0$ comme ensembles, puisque U_0 contient toutes les fonctions continues bornées. D'autre part, la topologie de U est plus fine que celle de U_0; en effet la topologie de U_0 n'est autre que la topologie de la convergence uniforme, et la topologie de U est cette même topologie uniforme augmentée de la condition que les fonctions qui tendent vers une limite et cette limite elle-même sont toutes sous une même fonction de U, c'est-à-dire dans un même U_{u_i}.

Théorème IV. Comme espace topologique, $M_0 \subseteq M$.

Cela résulte immédiatement de ce que M_0 et M sont les duals forts de U_0 et U (en prenant $X' = R$) et de ce que la topologie de U est plus fine que celle de U_0.

Remarque. Il résulte de ce théorème que, quand on a à faire à

des mesures finies, ce qui est le cas du calcul des probabilités, il est préférable d'utiliser la topologie M_0 qui est la meilleure possible dans le cadre choisi.

Théorème V

Les fonctions sommables (M) sont sommables (M_0); en particulier, les ensembles intégrables (M) sont intégrables (M_0).

Cela résulte de la définition des fonctions sommables ou des ensembles intégrables et des théorèmes III et IV ci-dessus.

II. Topologie M_0 et topologie de Prohorov

Pour montrer ce que signifie la topologie de M_0 quand on considère les mesures sous l'angle de fonctions complètement additives d'ensembles, nous allons donner une autre définition de la topologie dans le cas où l'espace topologique X est un espace *métrique*. Cette nouvelle définition sera une extension naturelle de celle adoptée par M. Prohorov [38] dans le cas de mesures de probabilités sur un espace R^n.

Soit B_0 le σ-anneau de parties contenant tous les ouverts de X.

Remarquons que B_0 est tout aussi bien le σ-anneau contenant tous les fermés de X : en effet, un fermé est la différence de l'espace entier et de son complémentaire qui sont tous deux ouverts, et inversement un ouvert est la différence de l'espace entier et de son complémentaire qui sont tous deux fermés.

Proposition I

Toute mesure de M_0 est définie sur un σ-anneau comprenant B_0. Cela résulte du théorème II.

Définition 2

Nous dirons qu'une suite (m_i) de mesures, définies sur un σ-anneau contenant B_0, converge au sens de Prohorov vers une mesure m_0, définie sur un σ-anneau comprenant aussi B_0, si pour tout A de B_0 tel que la mesure m_0 de sa frontière soit nulle, on a :

$$m_i(A) \to m_0(A).$$

Théorème VI

La convergence au sens de la topologie dans M_0 entraîne la convergence au sens de Prohorov.

Démonstration

Lemme. Soit F un fermé de mesure m_0 nulle : $m_0(F) = 0$.

Si $m_i \to m_0$ au sens de M_0, alors $m_i(F) \to 0$.

En effet, F est mesurable pour m_i puisque F est dans B_0.

Soit alors (f_k) une suite de fonctions de U_0 telles que :

$f_k(x) = 1$ sur F,

$f_k(x) = 0$ sur le complémentaire de F_ε, où F_ε est un voisinage arbitrairement petit de F,

$$\forall \varepsilon, \exists K : k \geqslant K \Rightarrow |m_0(f_k) - m_0(\varphi F)| \leqslant \varepsilon,$$

(où φF désigne la fonction caractéristique de F).

Une telle famille existe, puisque F est mesurable.

On a :

$$\begin{aligned} |m_i(F)| &\leqslant |m_i(f_k)| \leqslant |m_i(f_k) - m_0(f_k)| + |m_0(f_k)| \\ &\leqslant |m_i(f_k) - m_0(f_k) + |m_0(f_k) - m_0(\varphi F)| + m(\varphi F)|. \end{aligned}$$

Or $\forall \eta, \exists K : k \geqslant K \Rightarrow |m_0(f_k) - m_0(\varphi F)| \leqslant \eta$;

$\forall \eta, \exists I : i \geqslant I \Rightarrow |m_i(f_k) - m_0(f_k)| \leqslant \varepsilon$,

car $m_i \to m_0$ au sens de M_0;

$$m_0(\varphi F) = m_0(F) = 0.$$

Donc :

$$\forall \varepsilon, \exists I : i \geqslant I \Rightarrow m_i(F) \leqslant \varepsilon.$$

C. Q. F. D.

Nous pouvons maintenant passer à la démonstration du théorème. Fr A étant un fermé, on a, en vertu du lemme précédent,

$$m_i(\mathrm{Fr}\ A) \to 0,$$

puisque, par hypothèse, $m_0(\mathrm{Fr}\ A) = 0$.

On peut donc se limiter à l'étude de l'intérieur de A : Å.

Comme Å est un ouvert, il est intégrable; de plus sa mesure, qui est égale à sa mesure extérieure, est la limite sup des mesures des fonctions continues qui sont inférieures à sa caractéristique, du moins si on se limite aux parties positives des mesures

$$m^*_k(Å) = m_k(\varphi Å) = \sup_{\substack{g \in R_+ U_0 \\ g \leqslant \varphi Å}} m_k(g);$$

donc :

$$\forall \varepsilon, \exists k, \exists g_{k,\varepsilon} : m_k(g_{k,\varepsilon}) \geqslant m_k(\varphi Å) - \varepsilon.$$

Comme on a :

$$\begin{aligned}|m_k(\text{Å}) - m_0(\text{Å})| &= |m_k^*(\varphi\text{Å}) - m_0(\varphi\text{Å})| \\ &\leqslant |m_k^*(\varphi\text{Å}) - m_k(g_{k,\varepsilon})| + |m_k(g_{k,\varepsilon}) - m_0(g_{k,\varepsilon})| \\ &\quad + |m_0(g_{k,\varepsilon}) - m_0(\varphi\text{Å}),\end{aligned}$$

le théorème est démontré; car le premier terme tend vers 0 par définition des $g_{k,\varepsilon}$; le second terme tend vers 0 puisque les m_i convergent vers m_0 au sens de la topologie de M_0; enfin on peut évidemment prendre les $g_{k,\varepsilon}$ tels que le troisième terme tende vers 0.

La même chose peut être faite avec les parties négatives des mesures. Le théorème est donc démontré.

Théorème VII. La convergence au sens de Prohorov entraîne la convergence au sens de la topologie dans M_0.

C'est-à-dire que si, pour tout A de B_0 tel que $m_0(\text{Fr A}) = 0$, on a :

$$m_i(\text{A}) \to m_0(\text{A}),$$

alors $|m_i - m_0| \to 0$.

Démonstration. Nous nous limiterons aux mesures positives.

Lemme

Pour tout F fermé et pour tout $\varepsilon > 0$, il existe un $\eta > 0$ tel que le voisinage $F\eta$ d'ordre η de F ait sa frontière de mesure nulle pour m_0 et ait une mesure qui diffère de celle de F de moins de ε.

En effet,

1°) Soit (F_n) la suite des voisinages ouverts d'ordre $1/n$ de F. On a :

$$\bigcap_{n=1}^{\infty} F_n = F.$$

Or, on sait que, si (A_n) est une suite décroissante d'ensembles intégrables, on a

$$m\left(\bigcap A_n\right) = \lim m(A_n),$$

(cfr, par exemple, Halmos [16]). Donc, pour tout $\varepsilon > 0$, on peut trouver un voisinage ouvert d'ordre $1/N$ tel que :

$$m_0(F_N) < m_0(F) + \varepsilon.$$

2°) Considérons les voisinages fermés de F : $F\eta$ pour tous les η tels que

$$0 < \eta < 1/N.$$

Tous ces voisinages ont une mesure qui diffère de celle de F de moins de ε. Pour deux η_1, η_2 différents, ces voisinages ont des frontières disjointes puisque chacun de leurs points est situé à des distances différentes de F.

D'autre part l'ensemble des η tels que :

$$0 < \eta < 1/N$$

a la puissance du continu. On a donc une infinité non dénombrable d'ensembles disjoints, les Fr $F\eta$, tels que la mesure de leur réunion soit bornée. Il ne peut donc y avoir qu'une infinité dénombrable de termes de mesure non nulle et il existe donc un η tel que

$$m_0(\mathrm{Fr}\ F\eta) = 0.$$

C. Q. F. D.

Remarque. $\forall\ \zeta > 0$, on peut choisir le η du lemme précédent tel que : $\eta \leqslant \zeta$.

En effet, dans la démonstration précédente, au lieu de considérer l'ensemble des η tels que

$$0 < \eta < 1/N,$$

on peut considérer l'ensemble des η tels que

$$0 < \eta < \min\,[1/N,\ \zeta];$$

cet ensemble a encore la puissance du continu et la démonstration reste donc valable.

Démonstration du théorème VII

Il suffit de montrer que pour une fonction quelconque u de U_0, on a :

$$m_i(u) \to m_0(u).$$

A) u étant bornée, soit

$$0 \leqslant |u| \leqslant C.$$

Divisons l'intervalle [0, C] en n parties égales par les points

$$C_0 = 0,\ C_1, C_2, \ldots, C_i, \ldots, C_n = C.$$

Soit D_i l'ensemble des points de X tels que

$$C_i \leqslant |u(x)| \leqslant C_{i+1}.$$

Ces D_i sont fermés.

Pour tout $\varepsilon > 0$, on peut trouver un $\zeta > 0$ tel que sur un voisinage d'ordre ζ de D_i, la fonction diffère de moins de ε de ce qu'elle est sur D_i.

On a donc pour tout x de $D_{i,\zeta}$:

$$C_i - (C/n + \varepsilon) \leqslant |u(x)| \leqslant C_i + (C/n + \varepsilon), \qquad (1)$$

En vertu du lemme et de la remarque on peut alors trouver un

$$\eta_i \leqslant \zeta$$

tel que :

$$m_0 (\mathrm{Fr}\, D_{i,\eta_i}) = 0;$$

puisque $\eta_i \geqslant \zeta$, la relation (1) est a fortiori vérifiée pour tout x de D_{i,η_i}.

B) Nous avons donc obtenu un nombre fini de fermés D_{i,η_i} constituant un recouvrement de X, tous de frontières de mesure m_0 nulle, et sur lesquels u diffère peu d'une constante. On peut donc y associer une partition en un nombre fini d'ensembles :

$$E_k (n, \varepsilon), \quad 0 \leqslant k \leqslant l,$$

qui soient tous de frontière m_0 nulle et tels que, si a_k est une valeur de $|u(x)|$ sur E_k, on ait :

$$a_k - (C/n + \varepsilon) \leqslant |u(x)| \leqslant a_k + (C/n + \varepsilon)$$

pour tout x de E_k.

C) Si b_k est une valeur de u_k sur E_k, on a donc que les $g_{n,\varepsilon}$:

$$g_{n,\varepsilon} = \sum_{k=0}^{l} b_k \varphi E_k,$$

(où φE_k désigne la caractéristique de E_k), forment une suite de fonctions qui converge uniformément vers u, pour $n \to \infty$ et $\varepsilon \to 0$.

D) Or chacun des E_k est maintenant tel que :

$$m_i(E_k) \to m_0(E_k),$$

puisque, par hypothèse, les m_i convergent vers m_0 au sens de Prohorov.

On a donc :

$$m_i(g_{n,\varepsilon}) \to m_0(g_{n,\varepsilon}),$$

et, puisque les g convergent uniformément vers u,

$$m_i(u) \to m_0(u).$$

C. Q. F. D.

III. Mesures transformées

Le but de ce paragraphe est d'étudier les transformations de l'espace X métrique en un espace de *Banach* Y, qui transforment en un certain sens des mesures de $M_0(X)$ en mesures de $M_0(Y)$ et ce d'une manière compatible (en un sens à définir) avec les topologies M_0.

Soit $y = f(x)$ une transformation de l'espace X en un espace de Banach Y, transformation définie pour tout x de X.

Soit m une mesure de $M_0(X)$.

Définition 3

Nous définirons la mesure $f(m)$ de $M_0(Y)$, transformée de m par f, par sa valeur sur toute fonction $u(y)$ de $U_0(Y)$:

$$f(m)(u(y)) = m(u(f(x))),$$

si cette dernière expression a un sens.

Théorème VIII

Une condition suffisante pour que $f(m)$ soit bien définie comme mesure de M_0 (Y) est que $f(x)$ soit une fonction de X dans l'espace de Banach Y, intégrable pour la mesure m.

Démonstration

Dire que $m(u(f(x)))$ existe, c'est dire que $u(f(x))$ est sommable comme fonction de x pour la mesure m.

a) Montrons d'abord qu'elle est mesurable.

Puisque $f(x)$ est intégrable, elle peut être approchée par des fonctions dénombrablement valuées, et ce pour la topologie de la convergence en moyenne (m) :

$$f_n = \sum_{i=1}^{\infty} y_i \varphi A_i \to f(x), (m),$$

c'est-à-dire que :

$$m(f_n(x)) = m(\Sigma\, y_i\, m(A_i) \to m(f(x)), \qquad (1).$$

$u(f_n(x))$ s'obtient facilement puisque, pour un x donné, $f_n(x) = y_i$ pour une valeur fixée de i. On a :

$$u(f_n(x)) = \sum_i u(y_i)\ \varphi A_i.$$

$u(f_n(x))$ est donc dénombrablement valuée. Il reste à voir que $u(f_n(x))$ tend vers $u(f(x))$ pour la topologie de la convergence en moyenne (m). Pour cela, il suffit de voir que :

$$\sum_i u(y_i) m(A_i) \to m(u(f(x))),$$

ce qui résulte de la continuité de $u(x)$ et de (1). En effet :

i) En vertu de (1), pour tout $\varepsilon > 0$, on peut prendre n assez grand pour que, quel que soit le x fixé, on ait :

$$|y_i - f(x)| \leqslant \varepsilon.$$

ii) Or, en vertu de la continuité de $u(y)$, pour tout $\eta > 0$, on peut trouver un $\varepsilon > 0$ tel que :

$$|y_i - f(x)| \leqslant \varepsilon \Rightarrow |u(y_i) - u(f(x))| \leqslant \eta.$$

Donc, par définition de l'intégrale à la Bochner, on a bien :

$$\sum_i u(y_i)\ m(A_i) \to m(u(f(x))).$$

Donc $u\ (f(x))$ est mesurable.

b) $u(f(x))$ est alors intégrable; car $u(y)$ étant bornée, on a :

$$|u(f(x))| \leqslant M.\ |f(x)|;$$

$f(x)$ étant intégrable à la Bochner, $|f(x)|$ est intégrable à la Lebesgue. Donc $|u(f(x))|$, mesurable et inférieure à une fonction intégrable à la Lebesgue, est intégrable à la Lebesgue. Donc $u(f(x))$ est intégrable (m).

Théorème IX

Soit (m_i) une suite de mesures $M_0(X)$ où X est un espace métrique, convergeant vers une mesure m_0 de $M_0(X)$. Soit $y = f(x)$ une transformation partout définie de X en un espace de Banach Y, qui soit sommable pour la mesure m_0. Alors les transformées $f(m_i)$ des m_i par f existent dans $M_0(Y)$ à partir de i assez grand et les $f(m_i)$ convergent vers $f(m_0)$ dans $M_0(Y)$.

Démonstration

I) Montrons d'abord que les $f(m_i)$ existent pour i assez grand.

Il suffit, en vertu du théorème précédent, de montrer que $f(x)$ est sommable à la Bochner pour m_i à partir de i assez grand.

a) $f(x)$, étant intégrable (m_0), est limite pour la topologie de la convergence en moyenne (m_0) de fonctions intégrables, dénombrablement valuées : $f_n(x)$:

$$f_n(x) = \sum_i y_{i,n}\, \varphi A_{i,n}$$

$$m_0(f_n(x)) \to m_0(f(x)).$$

b) Je dis que, pour tout i et tout n, je peux choisir les $A_{i,n}$ de frontière de mesure nulle pour la mesure m_0.

i) Considérons pour un n donné la fermeture des $A_{i,n}$: $\overline{A}_{i,n}$ et considérons les fonctions :

$$f'_n(x) = \sum_i y'_{i,n}\, \varphi \overline{A}_{i,n}.$$

Pour un n donné, je peux choisir les $y'_{i,n}$ tels que :

$$m_0(f'_n(x)) = m_0(f_n(x)).$$

En effet, prenons :

$$y'_{1,n} = y_{1,n}.$$

Définissons $y'_{2,n}$ à partir de l'équation :

$$y_{2,n}\, m_0(A_2) = y'_{1,n}\, m_0\,(\overline{A}_1 \cap A_2) + y'_{2,n}\, m_0(\overline{A}_2) + y'_{2,n}\, m_0(\overline{A}_2 \cap A_1).$$

Et de même, par récurrence, définissons en général $y'_{i,n}$ par l'équation suivante, (où nous n'écrivons pas l'indice n),

$$y_i\, m_0(A_i) = \sum_{k=1}^{i-1} y'_k\, m_0(\overline{A}_k \cap A_i) + y'_i\, m_0(A_i) + \sum_{k=1}^{i-1} y'_i\, m_0(\overline{A}_i \cap A_k).$$

On vérifie aisément que, pour ces valeurs des $y'_{i,n}$, (qui existent puisqu'on est dans un espace de Banach Y), on a

$$m_0(f'_n(x)) = m_0(f_n(x)).$$

ii) Les $\overline{A}_{i,n}$ étant maintenant des fermés, on peut appliquer à chacun d'eux le lemme du théorème VII ; il existe donc pour tout $\varepsilon > 0$ un voisinage de $A_{i,n}$ dont la mesure m_0 diffère de celle de $A_{i,n}$ de moins de ε et qui ait sa frontière de mesure m_0 nulle.

iii) Nous pouvons alors prendre pour un n fixé une suite de ε_i telle que :

$$\sum_i y'_{i,n}\, \varepsilon_i$$

converge en norme vers 0. Cela est possible; en effet :

$$\sum_i y'_{i,n}\ m_0(\overline{A_{i,n}})$$

converge bien pour un n donné.

iv) Il suffit maintenant de passer de ce recouvrement à une partition de X, les ensembles de cette partition étant de frontière de mesure m_0 nulle. Faisons correspondre à chaque élément $B_{i,n}$ de cette partition l'élément $z_{i,n}$ de l'espace Y défini par :

$$z_{i,n} = \Sigma\, y'_{k,n},$$

la somme portant sur tous les $y'_{k,n}$ tels que $B_{i.n} \subseteq$ le voisinage considéré de $A_{k,n}$.

On a toujours que les fonctions :

$$\sum_i z_{i,n}\ \varphi\, B_{i,n}$$

sont, en mesure m_0, égales aux fonctions $f_n(x)$.

On a donc bien trouvé des ensembles de frontière de mesure m_0 nulle tels que :

$$g_n(x) = \sum_i z_{i,n}\ \varphi\, B_{i,n}$$

soient des fonctions qui convergent en moyenne (m_0) vers $f(x)$. Notons de nouveau ces fonctions $f_n(x) = \sum_i y_{i,n}\ \varphi\, A_{i,n}$.

c) Maintenant les $A_{i,n}$ ayant leur frontière de mesure m_0 nulle, on a :

$$m_j(A_{i,n}) \to m_0(A_{i,n}),$$

en vertu de l'équivalence de la topologie m_0 et de la topologie de Prohorov (Théorème VI). Ceci va nous permettre de montrer que, pour j assez grand, $m_j(f(x))$ existe.

En effet :

i) Puisque $m_0(\mathrm{Fr}\ A_i) = 0$:

$$\forall \varepsilon,\ \exists\ J : j \geqslant J \Rightarrow \left| m_j(A_{i,n}) - m_0(A_{i,n}) \right| \leqslant \varepsilon.$$

ii) Posons :

$$\varepsilon_i = \varepsilon\, m_0(A_{i,n}).$$

On a :

$$\left| \Sigma\ y_{i,n}\, m_j(A_{i,n}) - \sum_i y_{i,n}\, m_0(A_{i,n}) \right|$$

$$\leqslant \Sigma\ |y_{i,n}|\, [m_j(A_{i,n}) - m_0(A_{i,n})] \leqslant \varepsilon \sum_i |y_{i,n}|\, m_0(A_{i,n}).$$

iii) La mesure m_j et la mesure m_0 de f_n diffèrent donc d'aussi

peu que l'on veut; il en résulte que $f(x)$ est aussi bien la limite des $f_n(x)$ pour la topologie de la convergence en moyenne (m_i), avec i assez grand, que pour la topologie de la convergence en moyenne m_0.

II) Montrons maintenant que $f(m_i)$ converge vers $f(m_0)$ au sens de la topologie de $m_0(Y)$. C'est-à-dire que pour toute fonction $u(y)$ continue et bornée on a :

$$m\,(u(f(x))) \to m_0(u(f(x))).$$

En effet,

$$\begin{aligned}
&|m_i(y(f(x))) \quad - m_0(u(f(x)))| \\
\leqslant &|m_i(u(f(x))) \quad - m_i(u(f_n(x)))| \\
+ &|m_i(u(f_n(x))) - m_0(u(f_n(x)))| \\
+ &|m_0(u(f_n(x))) - m_0(u(f(x)))|.
\end{aligned}$$

Le premier et le troisième terme sont aussi petits que l'on veut; cela résulte de la démonstration du théorème VIII.

Le deuxième terme est aussi petit que l'on veut pourvu que i soit assez grand; en effet :

$$m_i\,(u(f_n(x))) = m_i(u(\sum_j y_{j,n}\varphi A_{j,n})) = m_i(\sum_j u(y_{j,n})\varphi A_{j,n}) = \sum_j u(y_{j,n}) m_i(\varphi A_{j,n}),$$

$$m_0(u_n(f(x))) = \sum_j u(y_{j,n})\; m_0(A_{j,n});$$

et $m_i(A_{j,n})$ tend bien vers $m_0(A_{j,n})$, puisque la frontière de $A_{j,n}$ est de mesure m_0 nulle.

Le théorème est donc démontré.

Chapitre III

MESURES FINIES SUR UN ESPACE DE BANACH : THÉORÈME CENTRAL LIMITE

Etant en possession de la théorie générale des mesures et des mesures finies, nous pouvons dans ce chapitre passer à la démonstration du théorème central limite pour des mesures sur un *espace de Banach*. Ce théorème a été démontré pour les espaces de Hilbert séparables dans la thèse de doctorat de Mlle Mourier [26] et pour des espaces un peu plus généraux par Fortet et Mourier [23].

§ 1. Notions préliminaires

A. *Topologie M_0 et Topologie M_b*

Parmi les espaces U qui peuvent servir à la définition des espaces de mesures, nous avons déjà signalé l'intérêt particulier de l'espace U_0 des fonctions continues et bornées. Il est intéressant de considérer ici également le cas où l'espace U est formé des fonctions continues bornées nulles en dehors d'une boule. Nous le désignerons par U_b.

Proposition I

U_b est bien un espace du type U, c'est-à-dire vérifie l'axiome A, (chapitre I).

En effet, il est immédiat que U_b est un espace vectoriel; U_b vérifie l'axiome A, car toute fonction inférieure à une fonction u de U_b est nulle en dehors de la même boule que celle en dehors de laquelle u est nulle.

De plus U_b est limite inductive d'une suite d'espaces de Banach, car on peut considérer une suite dénombrable de boules de centre 0 et de rayons croissants, par exemple la suite des boules B_n :

$$B_n = \{x : |x| \leqslant n\}.$$

Toute boule de X finit par être dans une des B_n pour n assez grand. D'autre part, pour toute boule B_n, il existe une fonction de U_b égale à 1 sur B_n; en effet il suffit de considérer la boule de rayon $n + \varepsilon$; il existe une fonction continue bornée nulle en dehors de $B_{n+\varepsilon}$ et égale à 1 sur B_n, puisque X, étant un espace de Banach, est métrique, donc normal.

On peut donc trouver une suite u_i de fonctions de U_b telles que pour toute fonction v de U_b il existe un I et des C_i avec :

$$i \geqslant I \Rightarrow |v(x)| \leqslant C_i |u_i(x)|.$$

Corollaire. Comme espace topologique on a

$$U_b \subseteq U_0,$$

$$M_0 \subseteq M_b.$$

C'est, en effet, un cas particulier des théorèmes III et IV du § I du chapitre précédent.

Le théorème V du même chapitre donne de même que les fonctions

sommables (M_b) sont sommables (M_0) et que les ensembles intégrables (M_b) sont intégrables (M_0).

Remarquons que U_b et M_b ne sont pas des espaces de Banach mais des limites inductives d'espaces de Banach.

B. *Moments et classes de mesure*

Soit $v(x)$ une fonction continue de X dans R. Soit m une mesure appartenant à un certain espace M de mesures (finies ou non).

Définition 1

Nous dirons que le v-moment de Slutsky existe pour une mesure m si :

$$v \in V(m)$$

Définition 2

Nous appellerons espace M_v l'espace des mesures de M telles que le v-moment de Slutsky existe pour chacune d'elle.

Définition 3

Nous désignerons par U_v l'espace vectoriel des fonctions continues de X dans X′ telles qu'il existe un C indépendant de x avec, pour tout x :

$$|u(x)| \leqslant C\,|v(x)|$$

Théorème I

U_v est formé de fonctions intégrables pour toute mesure de M_v.

C'est évident à partir des définitions. En effet, u étant continue est mesurable et $v(x)$ étant intégrable à la Bochner, $|v(x)|$ est est intégrable à la Lebesgue, donc $|u(x)|$ est intégrable à la Lebesgue, donc $u(x)$ est intégrable à la Bochner.

Remarque. Il est clair que si u est dans U_v, $U_u \subseteq U_v$; donc $M_v \subseteq M_u$.

Définition 4

Le v-moment de Slutsky est par définition :

$$m(v).$$

Définition 5

Nous appellerons v-moment-mesure de la mesure m la mesure $(v.m)$, si elle existe, définie, pour tout u de U, par :

$$(v.m)\,(u) = m(v.u).$$

Remarque. Si le v-moment de Slutsky existe, le v-moment-mesure existe aussi; puisque u étant bornée, $v.u$ est intégrable avec v. Mais la réciproque n'est pas vraie; le v-moment-mesure peut exister sans que le v-moment de Slutsky existe.

C'est le cas si et seulement si la mesure m n'est pas dans M_0. En effet, si m est dans M_0, le fait que le v-moment-mesure existe implique que $v.u$ soit intégrable pour tout u de U_0, en particulier pour $u = 1$; donc v doit être intégrable, c'est-à-dire que le v-moment de Slutsky existe.

Inversement, si m n'est pas dans M_0, il existe des v-moments-mesures tels que les v-moments de Slutsky correspondants n'existent pas. C'est, en effet, le cas notamment des v-moments pour $v = C^{te}$, puisqu'alors, $v.u$ se réduit à Cu qui est toujours intégrable quand u est dans U. Mais C n'est pas intégrable, sinon m serait dans M_0.

Exemple. Soit X un espace de Banach, U_b l'espace vectoriel des fonctions continues nulles en dehors d'une boule. Les fonctions

$$(< x^*, x >)^a \text{ ou } |x|^a, \qquad (a > 0),$$

sont bornées sur toute boule; les moments-mesures correspondants existent donc toujours mais les moments de Slutsky correspondants n'existent pas nécessairement.

Remarque. La mesure $v.m$ sera aussi désignée par l'expression « produit multiplicatif de m par v ».

Nous désignerons par M_a les mesures telles que les moments de Slutsky correspondant aux fonctions :

$$|x|^a$$

existent.

C. *Mesure caractéristique. Fonction caractéristique*

Définition 6

On appellera mesure caractéristique Fm correspondant à la mesure m, le produit multiplicatif :

$$Fm = e^{-i<x^*,x>}.\, m.$$

C'est une famille de mesures dépendant du paramètre x^*, variant dans le dual X^* de l'espace de Banach X.

Définition 7

On appellera fonction caractéristique correspondant à la mesure

m, la valeur de la mesure caractéristique appliquée à la fonction $u = 1$, si cette valeur existe, c'est-à-dire si la mesure caractéristique est dans M_0.

$$Fm(1) = m(e^{-i<x^*,x>}).$$

Remarque. On définit de même la mesure caractéristique inverse par le produit multiplicatif :

$$F^{-1}\, m = e^{+i<x^*,x>}.m$$

et la fonction caractéristique inverse par :

$$F^{-1}m(1) = m(e^{+i<x^*,x>}).$$

Théorème II

La connaissance de la mesure caractéristique ou de la fonction caractéristique détermine complètement la mesure m comme mesure d'un espace M.

Démonstration

I) La connaissance de la mesure caractéristique détermine la mesure; en effet, Fm et $F^{-1}m$ existent toujours et sont dans M avec m, puisque :

$$|e^{+i<x^*,\,x>}| = 1.$$

Donc les mesures :

$$FF^{-1}m \text{ et } F^{-1}Fm$$

existent toujours également dans M et on a évidemment :

$$FF^{-1}m = F^{-1}Fm = m$$

II) m, comme mesure de M, est aussi déterminée par sa fonction caractéristique. En effet,

i) Comme l'ensemble des $< x^*, x >$ sépare les points de X, il en est de même des fonctions :

$$a\, e^{-i\,<x^*,\,x>}$$

où x^* parcourt X^*.

ii) En vertu du théorème de Stone-Weierstrass généralisé, (Ch. I, § 5), toute fonction sommable peut alors être approchée, si elle est bornée en mesure, par des polynomes par rapport aux fonctions

$$a\, e^{-i<x^*,\,x>}.$$

iii) Chaque terme de ces polynomes est du type :

$$C\, e^{-ki<x^*,\,x>}$$

donc de norme bornée; ces termes sont donc intégrables si $a\, e^{-i<x^*, x>}$ l'est, c'est-à-dire si la fonction caractéristique existe.

iv) L'ensemble des polynomes en $a\, e^{-i<x^*, x>}$ étant donc dense dans U pour la topologie de la convergence en moyenne et ces polynomes étant intégrables, la mesure est bien déterminée sur U par la fonction caractéristique si elle existe.

Théorème III

Si (m_i) est une suite de mesures convergeant vers une mesure m_0, alors les mesures Fm_i convergent et convergent précisément vers Fm_0. Réciproquement, si une suite de mesures transformées : Fm_i converge vers une mesure limite $n(x^*)$, dépendant du paramètre x^*, les m_i convergent vers une mesure m_0 et :

$$n(x^*) = Fm_0.$$

Démonstration

L'espace M des mesures étant limite inductive d'espaces de Banach, il suffit de montrer que sur tout U_u, les normes d'une mesure et de sa mesure caractéristique sont les mêmes.

1°)
$$\begin{aligned} |Fm|_u &= \sup_{v \in U_u} |Fm(v)| \\ &= \sup_{v \in U_u} |m\ (e^{-i<x^*, x>}\, v(x)|, \end{aligned} \qquad (1).$$

D'autre part :

$$|m|_u = \sup_{v \in U_u} m(v) \qquad (2).$$

Comme :

$$|e^{-i<x^*, x>}.v| = |v|,$$

on ne prend dans (1) pour chercher la borne supérieure qu'une partie des v que l'on prend dans (2); la borne supérieure ne peut donc qu'être plus petite, et :

$$|Fm|_u \leqslant |m|_u.$$

2°) Inversement, (2) peut s'écrire :

$$|m|_u = \sup_{v \in U_u} |e^{-i<x^*, x>}\, e^{+i<x^*, x>}\, v|, \qquad (3);$$

car : $$|e^{-i<x^*, x>}| = |e^{+i<x^*, x>}| = 1.$$

On ne prend donc dans (3) qu'une partie des v que l'on prend dans (1). On a donc :

$$|Fm|_u \geqslant |m|_u.$$

Donc :

$$|Fm|_u = |m|_u ;$$

d'où la thèse.

Remarque. On peut évidemment partout remplacer F par F^{-1}.

Remarque. Le théorème précédent n'est pas applicable aux fonctions caractéristiques. Un contre-exemple dû à Mlle Mourier le montre dans le cas d'un espace X de Hilbert [26].

D) *Développement d'une mesure caractéristique*

Le développement d'une mesure caractéristique suivant ses moments mesures peut être poussé plus ou moins loin suivant la topologie que l'on envisage. Nous examinerons les deux cas suivants : topologie dans M_b (dual des fonctions continues nulles en dehors d'une boule) et topologie dans M_0 (dual des fonctions continues et bornées).

Théorème IV

Pour toute mesure m de M_b et pour tout s, on a :

$$Fm = \sum_{n=1}^{\infty} \frac{(-i < x^*, x >)_n}{n\,!} \cdot m$$

$$= m + (-i) < x^*, x > .m + (-i)^2 \frac{< x^*, x >^2}{2} \cdot m$$

$$+ \ldots + (-i)^s \frac{< x^*, x >^s}{s\,!} .m + < x^*, x >^s .w\,(x^*, x).m, \quad (1)$$

où $w(x^*, x).m$ converge vers 0 au sens de la topologie de M_b si $|x^*|$ tend vers 0.

Démonstration

Remarquons d'abord que le produit multiplicatif $f.m$ est linéaire en f et que les mesures-moments sont tous dans M_b puisque les fonctions $C < x^*, x >^k$ sont bornées sur toute boule. La formule (1) en résulte immédiatement. Il reste à montrer que $w(x^*, x).m$ tend vers 0 au sens de M_b.

Puisque tous les moments-mesures existent, (c'est-à-dire sont dans M_b), on peut écrire :

$$m = \sum_{n=1}^{\infty} \frac{(-i < x^*, x >)^n}{n!} . m = \sum_{n=1}^{\infty} \frac{(-i < x^*, x >)^n}{n!} . m);$$

donc :

$$w(x^*, x) . m = \sum_{n=s+1}^{\infty} \frac{(-i < x^*, x >)^{n-s} (-i)^s}{n!} . m),$$

et ce terme appliqué à une fonction de U_b est bien de l'ordre de $|x^*|$, puisque

$$|< x^*, x >| \leqslant |x^*| |x|$$

et que les fonctions de u ne sont différentes de θ que dans une boule, donc pour des $|x|$ bornés.

$w(x^*, x) . m$ converge donc bien vers θ au sens de la topologie de M_b.

Théorème V

Pour toute mesure de classe M_a $(a > \theta)$, où s est un entier contenu dans a, on a :

$$Fm = \sum_{n=1}^{\infty} \frac{(-i < x^*, x >)^n}{n!} . m$$

$$= m + (-i) < x^*, x > . m + (-i)^2 \frac{< x^*, x >^2}{2!} . m$$

$$+ \ldots + (-i)^s \frac{< x^*, x >^s}{s!} . m + < x^*, x >^s . w(x^*, x,) . m,$$

où $w(x^*, x) . m$ tend vers θ au sens de la topologie de M_0 si $|x^*|$ tend vers θ.

Démonstration

On vient de voir que $w(x^*, x) . m$ tend bien vers θ quand on l'applique à une fonction u de U_b : $(w(x^*, x) . m)(u)$ tend vers θ pour tout u de U_b. En effet, si m est de classe M_a, m est à fortiori dans M_b.

D'autre part, si on applique Fm à une fonction $u = 1$, on obtient la fonction caractéristique. Or, pour la fonction caractéristique, Mlle E. Mourier, [36 p. 49], a montré le résultat ci-dessus, à savoir le développement limité. Donc : $(w(x^*, x) . m$ (C^{te}) tend vers θ. Comme l'ensemble $\{U_b \cup C^{tes}\}$ engendre un espace vectoriel dense dans l'espace U_0 des fonctions continues et bornées pour la

topologie de U_0, $(w(x^*, x).m)(u)$ tend vers 0 pour tout u de U_0. Donc $w(x^*, x).m$ tend bien vers 0 au sens de la topologie de M_0.

E. *Mesures indépendantes.*

Soit l une mesure sur le produit d'un espace de Banach X par lui-même. Soit m_1 et m_2 les mesures définies de la manière suivante :

$$m_1(A) = l(A \times X),$$
$$m_2(A) = l(X \times A).$$

C'est un fait bien connu que m_1 et m_2 sont des mesures définies sur le σ-anneau de parties des ensembles de X tels que $A \times X$ et $X \times A$ soient mesurables pour la mesure l.

Définition 8

Nous dirons que l est le produit direct de deux mesures indépendantes m_1 et m_2 si on a, pour tout A et tout B mesurables :

$$l(A \times B) = m_1(A) \times m_2(B).$$

Nous utiliserons le théorème suivant, (cfr. Halmos, [16]) : Pour toute fonction réelle $u(x, y)$, sommable (1), telle que :

$$u(x, y) = f(x).g(y)$$

où f et g sont sommables, on a :

$$l(u(x, y)) = m_1(f(x)) \cdot m_2(g(y)).$$

Théorème VI

Soit $W_{x,y}$ les fonctions continues et bornées à valeurs réelles telles que :

$$u(x, y) = f(x).g(y).$$

Je dis que la connaissance de l sur $W_{x,y}$ détermine complètement l comme mesure de type M_0 sur $X \times X$.

Démonstration

Il suffit de montrer que l'ensemble $U_0(x, y)$ des fonctions continues et bornées sur $X \times X$ est dans la fermeture de $W_{x,y}$ pour la topologie de la convergence en moyenne (l). Comme d'autre part $U_0(x, y)$ est dans la fermeture, pour la même topologie, de l'espace vectoriel engendré par les fonctions caractéristiques mesurables, il suffit de montrer que ces fonctions caractéristiques sont dans

$\overline{W}_{x,y}$, (la barre indiquant la fermeture pour la topologie de la convergence en moyenne l).

Mais ceci est immédiat puisqu'aussi bien on a :

$$v(x, y) = v_1(x) . v_2(y)$$

si $v(x, y)$ est la fonction caractéristique de $A \times B$ et $v_1(x)$, $v_2(y)$ les fonctions caractéristiques respectives de A et B.

Remarque. En considérant comme $W_{x,y}$ non pas des fonctions réelles mais des fonctions à valeurs dans un espace de Banach X' telles que, pour tout x^* de X'*, on ait :

$$x^*u(x, y) = f(x^*, x) . g(x^*, y)$$

on a, comme dans le théorème ci-dessus, que l est déterminée sur $U_0(x, y)$ par sa connaissance sur $W_{x,y}$. En effet, soit

$$\begin{aligned} v(x, y) &= a \text{ si } (x, y) \in A \times B, \\ &= 0 \text{ si } (x, y) \notin A \times B; \\ v_1(x) &= a \text{ si } x \in A, \\ &= 0 \text{ si } x \notin A; \\ v_2(y) &= a \text{ si } y \in B, \\ &= 0 \text{ si } y \notin B. \end{aligned}$$

On a :

$$x^*(v(x, y)) = \varepsilon \sqrt{|x^*v_1(x)|} \cdot \sqrt{|x^*v_2(y)|}$$

avec $\varepsilon = +1$ si $x(a) \geqslant 0$,
$= -1$ si $x(a) \leqslant 0$.

Théorème VII

Si l est le produit direct de deux mesures indépendantes m_1 et m_2, on a, pour toute fonction $u(x, y)$ à valeur dans X' et appartenant à $W_{x,y}$ et pour tout z^* de X'* :

$$Fl(< z^*, u(x, y) >) = Fm_1(f(z^*, x)) . Fm_2(g(z^*, y)).$$

Démonstration

Par définition :

$$Fl(< z^*, u(x, y) >) = l(e^{-i<x^*, (x,y)>} < z^*, u(x, y) >)$$

avec $x^* \in (X \times X)^*$.

Or on sait que, (cfr. E. Mourier [4] p. 67), toute fonctionnelle linéaire sur $X \times X$ est du type :

$$<x^*, (x, y)> = <x^*_1, x> + <x^*_2, y>$$

avec : $x^*_1 \in X^*$ et $x^*_2 \in X^*$.

Donc, puisque $u(x, y)$ est dans $W_{x,y}$, on a :

$$Fl(<z^*, u(x, y)>) = l(e^{-i<x^*_1, x>}\, e^{-i<x^*_2, y>}\, f(z^*, x).\, g(z^*, y)).$$

Or m_1 et m_2 sont indépendantes; on a donc :

$$\begin{aligned} Fl(<z^*, u(x,y)>) &= m_1(e^{-i<x^*_1, x>} f(z^*, x)).\ m_2(e^{-i<x^*_2, y>} g(z^*, y)) \\ &= Fm_1(f(z^*, x)).\ Fm_2(g(z^*, y)). \end{aligned}$$

C. Q. F. D.

F. *Mesures laplaciennes.*

Définition 9.

Conformément à une définition de M. Fréchet [15], nous dirons qu'une mesure sur un espace de Banach X est laplacienne si toutes les transformées de cette mesure par les transformations

$$x \to < x^*, x >$$

pour tous les x^* de X^* sont des mesures laplaciennes sur R.

Fonction caractéristique de mesure laplacienne.

Mlle E. Mourier a montré quelle est la forme de la fonction caractéristique d'une mesure laplacienne sur un espace de Banach. En supposant les premiers moments nuls, c'est-à-dire que :

$$m(< x^*, x >) = 0,$$

son expression est dans nos notations :

$$Fm(1) = e^{-\frac{m(<x^*, x>^2)}{2}}.$$

§ 2. Le théorème central limite.

a) Composition de deux mesures.

Nous utiliserons la définition du produit de composition de deux mesures introduite par M. A. Weil [34] et adoptée par M. Schwartz

pour les distributions :

$$m_1 * m_2 (u) = m_1 \times m_2 (u(\xi_1 + \xi_2)).$$

Cette formule est valable si on considère des u à valeurs réelles et a un sens moyennant certaines conditions sur m_1 et m_2 (par exemple m_1 quelconque et m_2 à support compact). Les conditions les plus générales ne nous intéresseront pas ici puisqu'aussi bien il nous suffit de savoir que pour deux mesures de M_0 la définition a toujours un sens.

Si les fonctions u sont à valeurs dans un espace de Banach X', on peut de même conformément à la définition de l'intégrale de Pettis [77] écrire :

$$\forall z^* \in X^* \quad z^*(m_1 * m_2(u)) = m_1 * m_2 (z^* u) = m_1 \times m_2 (z^* u(\xi_1 + \xi_2)).$$

La même chose peut évidemment être faite avec n mesures de M_0 :

$$z^*.m_1 * m_2 * m_3 * \ldots * m_n(u) = m_1 \times m_2 \times m_3 \times \ldots \times m_n (z^* u(\xi_1 + \xi_2 + \ldots + \xi_n).$$

b) Soit W l'espace vectoriel des fonctions intégrables, engendré par les fonctions telles que, $\forall n \in N$, $\forall z^* \in X^*$:

$$z^* u(\xi_1 + \xi_2 + \ldots + \xi_n) = f(z^*, \xi_1).f(z^*, \xi_2)\ldots f(z^*, \xi_n).$$

Théorème VIII.

La connaissance de m sur W détermine complètement m.

Démonstration

Cela résulte de l'extension du théorème de Stone-Weierstrass. En effet,

i) On a dans W les fonctions du type :

$$e^{ib<x^*, x>}$$

et ce, $\forall b \in R$, $\forall x^* \in X^*$.

ii) Tout polynome en ces fonctions est intégrable, puisque chacun des termes de ces polynomes est encore de ce type, donc dans W.

iii) Ces fonctions séparent les points de X.

On est donc bien dans les conditions d'application du théorème généralisé de Stone-Weierstrass et, par continuité, la détermination de m sur W entraîne sa détermination sur U_0.

Théorème IX.

Si m_1 et m_2 sont indépendantes et si u est dans le système générateur de W, on a :

$$z^* \, \mathrm{F}m_1 \times m_2(u) = \mathrm{F}m_1(f(z^*, \xi_1) \cdot \mathrm{F}m_2(f(z^*, \xi_2)).$$

En effet,

$$\begin{aligned} z^* \, \mathrm{F}m_1 * m_2 \, (u) &= z^* \, e^{-i<x^*, \xi_1+\xi_2>} \cdot m_1 \times m_2 \, u(\xi_1 + \xi_2) \\ &= e^{-i<x^*, \xi_1>} \, e^{-i<x^*, \xi_2>} \cdot m_1 \times m_2 \, z^* u(\xi_1 + \xi_2) \\ &= (e^{-i<x^*, \xi_1>} \cdot m_1) \, (f(z^*, \xi_1)). \\ &\qquad\qquad (e^{-i<x^*, \xi_2>} \cdot m_2) \, (f(z^*, \xi_2)) \\ &= \mathrm{F}m_1 \, (f(z^*, \xi_1)). \, \mathrm{F}m_2(f(z^*, \xi_2)). \end{aligned}$$

La même chose est valable pour n mesures indépendantes. Si de plus ces mesures sont identiques, on a :

$$z^* \, \mathrm{F}m^{n*}(u) = [\mathrm{F}m(f(z^*, \xi))]^n.$$

Théorème central limite X.

La suite des mesures :

$$m^{n*}(u(x/\sqrt{n})),$$

où $m(u)$ est une mesure positive de classe M_2, converge au sens de la topologie de M_0 vers une mesure laplacienne.

Démonstration

a) Pour démontrer le théorème, il suffit de démontrer que les mesures caractéristiques convergent au sens de la topologie de M_0 vers une mesure caractéristique de mesure laplacienne; en effet, le théorème (III) donne alors la thèse.

b) Je dis que, pour démontrer que les mesures caractéristiques convergent, il suffit de montrer que, pour tout u de W, ces mesures convergent.

En effet :

Soit v une fonction quelconque de U_0 (continue et bornée).

Notons m_p la mesure $m^p(u(x/p))$.

En vertu du théorème de Stone-Weierstrass, il existe une suite (u_i) de fonctions de W, convergeant presqu'uniformément vers v. « Presque » signifie sauf pour un ensemble de mesure m nulle. Mais, comme les ensembles de mesure m nulle sont aussi des ensembles de mesure (m_p) nulle, puisque m_p n'est que la p^{me} composée

de m avec elle-même, la convergence est presque uniforme par rapport à m_p, indépendamment de p. Par conséquent les u_i convergent vers v en moyenne (m_p), indépendamment de p.

Pour que les mesures m_p convergent pour tout u de U, il suffit donc de vérifier qu'elles convergent pour tout u de W, car on a l'inégalité :

$$|m_p(v)-m_q(v)|\leqslant|m_p(v)-m_p(u_i)|+|m_p(u_i)-m_q(u_i)|+|m_q(u_i)-m_q(v)|.$$

Chacun des trois termes du second membre est arbitrairement petit, pourvu que l'on prenne p, q et i assez grands. Pour le premier et le troisième terme, cela résulte de la convergence presque uniforme des u_i vers v. Pour le deuxième terme, cela résulte de ce que les u_i sont dans W et que pour W, on a la convergence par hypothèse.

c) Puisque les mesures caractéristiques convergent pour la topologie M_0 lorsqu'on a vérifié qu'elles convergent pour tout u de W, il suffira alors de démontrer ce point, car l'espace M_0 étant complet, la convergence entraîne la convergence vers une mesure.

d) Vérifions maintenant la convergence des mesures caractéristiques pour une fonction u de W.

On a :

$$z^*\mathrm{F}m''^*(u(x/\sqrt{n})) = [\mathrm{F}m(f(z^*,\xi/\sqrt{n}))]^n$$
$$= [(e^{-i\langle x^*,x/\sqrt{n}\rangle}.m\,(f(z^*,x/\sqrt{n}))]^n$$
$$= \Big[\Big(1+\frac{(-i\langle x^*,x\rangle)}{\sqrt{n}}+\frac{(-1\langle x^*,x\rangle)^2}{2n}$$
$$+\frac{\langle x^*,x\rangle^2\,w(x^*/\sqrt{n})}{n}\Big).m\,(f(z^*,x/\sqrt{n})\Big]^n,$$

où $w(x^*)m$ tend vers 0 avec $|x^*|$ au sens de la topologie M_0.

Ensuite,

$$z^*\mathrm{F}m''^*(u(x\sqrt{n})) = \Big[m(f(z^*,x/\sqrt{n}))+\Big(\frac{(-i\langle x^*,x\rangle)}{\sqrt{n}}.m\Big)(f(z^*,x\sqrt{n}))$$
$$+\Big(\frac{(-i\langle x^*,x\rangle)^2}{2n}.m\Big)\,(f(z^*,x/\sqrt{n})$$
$$+\frac{\langle x^*,x\rangle^2 w\,(x^*/\sqrt{n})\,m\,(f(z^*,x/\sqrt{n}))}{n}\Big]^n.$$

$$= m(f(z^*, x/\sqrt{n}))^n \Big[1 + \frac{((-i <x^*, x>)^2 . m)(f(z^*, x/\sqrt{n}))}{n\, m(f(z^*, x/\sqrt{n}))} + \frac{(-i <x^*, x> . m)\,(f(z^*, x/\sqrt{n}))}{\sqrt{n}\, m(f(z^*, x/\sqrt{n}))} + \frac{(<x^*, x>^2 w(x^*/\sqrt{n})m)\,(f(z^*, x/\sqrt{n}))}{n\, m(f(z^*, x/\sqrt{n}))} \Big]^n$$

En prenant le log, (c'est évidemment un log complexe; on choisira donc une détermination, par exemple celle qui est continue et qui s'annule pour la fonction $u = 0$.)

$$\log Fm^{n*}(z^* u(x/\sqrt{n})) = n \log m(f(z^*, x/\sqrt{n})) + n \log (\ldots).$$

Nous écrirons plus brièvement :

$$\log A = n \log B + n \log \Big(1 + \frac{C + D + E}{n}\Big).$$

Or on sait que :

$$\log (1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \ldots,$$

où la série converge pourvu que $|z| \leqslant 1$.

Or nous allons voir que l'on peut prendre n assez grand pour que :

$$\left| \frac{C + D + E}{n} \right| \leqslant 1.$$

En effet, transformons le terme C :

$$C = \frac{((-i <x^*, x>)^2 . m)\,(f(z^*, x/\sqrt{n}))}{2\, m(f(z^*, x/\sqrt{n}))}.$$

Multiplions haut et bas par :

$$[m(f(z^*, x/\sqrt{n}))]^{n-1}.$$

$$C = \frac{((-i <x,^* x>)^2 . m)\,(f(z^*, x/\sqrt{n}))\; m(f(z^*, x/\sqrt{n}))^{n-1}}{2\, m(f(z^*, x/\sqrt{n}))_n}$$

que nous posons égal à

$$- \frac{F}{2m^{n*}(u(x/\sqrt{n}))}.$$

Réécrivons F en distinguant les n variables et en notant f_i pour :

$$f(z^*, \xi_i/\sqrt{n}).$$

On a :

$$F = \int <x^*, \xi_1>^2 f_1\, dm_1 \int f_2\, dm_2 \int \dots \int f_n\, dm_n;$$

donc, puisque les mesures sont indépendantes,

$$F = \int\int \dots \int <x^*, \xi_1>^2 \prod_{i=1}^{n} f_i\, dm_1\, dm_2 \dots dm_n;$$

donc aussi, puisque les mesures sont identiques :

$$F = \int\int \dots \int < x^*, \xi >^2 \prod_{i=1}^{n} f_i\, dm_1\, dm_2 \dots dm_n$$

$$= \int\int \dots \int \frac{\sum_{r=1}^{n} < x^*, \xi_r >^2}{n} \prod_{i=1}^{n} f_i\, dm_1\, dm_2 \dots dm_n.$$

Or,

$$(\sum_{r=1}^{n} < x^*, \xi_r >)^2 = \sum_{r=1}^{n} < x^*, \xi_r >^2 + 2 \sum_{\substack{p,q=1 \\ p<q}}^{n} < x^*, \xi_p >< x^*, \xi_q >.$$

Posant F = G — H, on a :

$$G = \int\int \dots \int \frac{(\sum_{r=1}^{n} < x^*, \xi_r >)^2}{n} \prod_{i=1}^{n} f_i\, dm_1\, dm_2 \dots dm_n,$$

$$G = \int\int \dots \int < x^*, \frac{\xi_1 + \xi_2 + \dots + \xi_n}{\sqrt{n}} >^2 \prod_{i=1}^{n} f_i\, dm_1\, dm_2 \dots dm_n;$$

ce qui, par définition de m^{n*}, donne :

$$G = m^{n*} (< x^*, x/\sqrt{n} >^2 z^* u(x/\sqrt{n}))$$

$$= (< x^*, x/\sqrt{n} >^2 \cdot m^{n*}) (z^* u(x/\sqrt{n})).$$

De même, le terme H se calcule comme suit :

$$H = 2 \sum_{\substack{p,q=1 \\ p<q}}^{n} \int\int \dots \int \frac{< x^*, \xi_p >< x^*, \xi_q >}{n} \prod_{i=1}^{n} f_i \cdot dm_i.$$

Comme les mesures sont indépendantes, on peut séparer les termes relatifs aux différentes mesures. Observons aussi que tous les termes de la somme sont égaux entre eux, puisque toutes les mesures sont égales. On a donc successivement :

$$H = 2\,\frac{n\,(n-1)}{2}\int\int\cdots\int \frac{<x^*,\,\xi_1><x^*,\,\xi_2>}{n}\prod_{i=1}^{n} f_i\,dm_i$$

$$= (n-1)\int <x^*,\,\xi_1 f_1\,dm_1\int <x^*,\,\xi_2> f_2\,dm_2\prod_{i=3}^{n}\int f_i\,dm_i$$

$$= n-1\,\frac{\left[\int <x^*,\xi_1> f_1\,dm_1\prod_{i\neq 1} f_i\,dm_i\right]\left[\int <x^*,\xi_2> f_2 d\,m_2\prod_{i\neq 2} f_i\,dm_i\right]}{\prod_{i=1}^{n}\int f_i\,dm_i}.$$

Dans chacun des crochets, on peut alors observer que, les mesures étant identiques, on peut permuter l'ordre des variables et réintroduire l'intégrale multiple, puisque les mesures sont indépendantes. Il vient pour tout k et tout l :

$$= \frac{(n-1)\left[\int\int\cdots\int <x^*,\xi_k>\prod_{i=1}^{n} f_i\,dm_i\right]\left[\int\int\cdots\int <x^*,\xi_1>\prod_{i=1}^{n} f_i\,dm_i\right]}{\prod_{i=1}^{n}\int f_i\,dm_i}.$$

En prenant dans chaque crochet la moyenne arithmétique des n termes égaux :

$$= \frac{(n-1)\left]\int\int\cdots\int \frac{<x^*,\sum_{r=1}^{n}\xi_r>}{n}\prod_{i=1}^{n} f_i\,dm\right]\left[\int\int\cdots\int \frac{<x^*,\sum_{r=1}^{n}\xi_r>}{n}\prod_{i=1}^{n} f_i\,dm_i\right]}{\prod_{i=1}^{n}\int f_i\,dm_i}$$

$$= (n-1)\,\frac{[(<x^*,\,x/\sqrt{n}>.m^{n*})\,(z^*u(x/\sqrt{n}))]^2}{m^{n*}(z^*u(x/\sqrt{n}))}.$$

D'où le terme C devient :

$$C = \frac{-(<x^*,\,x/\sqrt{n}>.m^{n*})\,(z^*u(x/\sqrt{n}))}{2m^{n*}(u(x/\sqrt{n})}$$

$$+\frac{(n-1)}{2n}\,\frac{[(<x^*,\,x/\sqrt{n}>.m^{n*})\,(z^*u(x/\sqrt{n}))]^2}{[m^{n*}(z^*u(x/\sqrt{n}))]^2}.$$

Calculons maintenant le terme D :

$$\frac{D}{n} = \frac{\dfrac{(-i < x^*, x >)^2}{\sqrt{n}} . m)\,(f(z^*, x/\sqrt{n}))}{m(f(z^*, x/\sqrt{n})} ;$$

en multipliant, comme plus haut, par $m(f(z^*, x/\sqrt{n}))^{n-1}$:

$$\frac{D}{n} = \frac{[(-i<x^*, x> . m)\,(f(z^*, x/\sqrt{n})\; m(f(z^*, x/\sqrt{n})]^{n-1}}{\sqrt{n}\, m(f(z^*, x/\sqrt{n})^n} .$$

D'où en distinguant le rôle des variables, puis en observant que l'on peut les échanger donc prendre la moyenne arithmétique de termes égaux, on obtient :

$$\frac{D}{n} = \frac{-i \int\int \dots \int \dfrac{\sum_{r=1}^{n} <x^*, \xi_i>}{\sqrt{n}} \prod_{i=1}^{n} f_i\, dm_i}{\sqrt{n}\, m^{n*}(z^* u(x/\sqrt{n}))}$$

$$= \frac{(-i <x^*, x/\sqrt{n}> . m^{n*})\,(z^* u(x/\sqrt{n}))}{n\, m^{n*}(z^* u(x/\sqrt{n})} .$$

Ceci établi nous pouvons utiliser la formule de développement du log, car les termes C, D, E sont bornés.

Les numérateurs sont en effet bornés sur toute fonction bornée puisque la mesure m est de classe M_2 et on peut se limiter à des fonctions telles que les dénominateurs soient supérieurs à un ϵ donné, puisque, pour calculer la norme dans M_0, on prend une borne sup et que les fonctions de mesure petite n'interviennent donc pas.

On a donc :

$$\log A = n \log B + \frac{C + D + E}{n} - \frac{1}{2} \cdot \left(\frac{C + D + E}{n}\right) + \dots ;$$

ce qui donne :

$$\log F m^{n*}(z^* u(x/\sqrt{n})) = \log m^{n*}(z^* u(x/\sqrt{n})) +$$

$$+ \frac{(-i <x^*, x/\sqrt{n}> . m^{n*})\,(z^* u(x/\sqrt{n}))}{m^{n*}(z^* u(x/\sqrt{n}))}$$

(I)
$$- \frac{(<x^*, /\sqrt{n}>^2 . m^{n*})\,(z^* u(x/\sqrt{n}))}{2\, m^{n*}(z^* u(x/\sqrt{n}))}$$

$$+\frac{(n-1)}{2n}\,\frac{(<x^*,\,x/\sqrt{n}>.m^{n*})\,(z^*u(x/\sqrt{n}))^2}{m^{n*}(z^*u(x/\sqrt{n}))^2}$$

$$+\,R,$$

où R tend vers 0 lorsque n tend vers l'infini, au sens de la topologie dans M_0.

e) Il reste maintenant à voir que la mesure limite, dont nous savons l'existence, est une mesure laplacienne.

Il suffit pour cela de faire $u = 1$ dans l'expression du log de la mesure caractéristique pour obtenir le log de la fonction caractéristique. Et ici, comme on n'a plus des moments-mesures mais des moments simples, on peut sans restreindre la généralité supposer que le premier moment est nul. On obtient pour la limite :

$$\log F\, m(1) = \log m(1) - \frac{(<x^*,\,x.>^2 m)\,(1)}{2\,m\,(1)}$$

ce qui est bien l'expression de la fonction caractéristique d'une mesure laplacienne.

C. Q. F. D.

Remarque : Théorème central limite local.

Le principe de la démonstration précédente suggère l'extension suivante du théorème central limite : si, au lieu de considérer la topologie M_0, de la convergence en loi on considère une autre topologie par exemple la topologie M_b : dual de l'ensemble des fonctions contenues nulles en dehors d'une boule, que peut-on dire de la limite de la suite des mesures :

$$m^{n*}(u(x/\sqrt{n}))\ ?$$

On vérifie aisément que toute la démonstration subsiste, jusque et y compris le développement du log de la mesure caractéristique. Mais cette limite ne peut plus être dite laplacienne sans abus de langage.

Définition 10

Une mesure sera dite localement laplacienne, si, pour toute fonction de U, (et non plus de U_0), on a l'expression (I) du log de la mesure caractéristique.

Théorème central limite local XI. La suite des mesures

$$m^{n*}(u(x/\sqrt{n}))$$

converge vers une mesure localement laplacienne au sens de la topologie dans M sous la seule réserve que le produit de composition existe.

Le produit de composition existe en particulier pour $m \in M_0$ sans restriction sur les moments.

Chapitre IV

CONVERGENCE DE LA RÉPARTITION EMPIRIQUE VERS LA RÉPARTITION THÉORIQUE

Le problème statistique de la convergence de la répartition empirique vers la répartition théorique a été posé et résolu dans le cas général d'éléments aléatoires à valeurs dans un espace métrique séparable par M. R. Fortet et Mlle E. Mourier [13].

Les méthodes développées dans ce travail vont permettre de préciser cette convergence et d'obtenir ainsi des résultats qui n'étaient connus que pour la droite réelle. Même dans le cas de l'espace R^n ($n > 1$), on n'avait en effet pas de renseignements sur la rapidité de la convergence.

Nous exposerons d'abord la position du problème adoptée par M. R. Fortet et Mlle E. Mourier. Nous aborderons ensuite deux méthodes pour étudier cette convergence. La première utilise un seul espace pour tous les problèmes mais ne permet pas de pousser les calculs sans estimations. La seconde utilise un espace variant avec chaque problème mais donne un résultat explicite général.

Comme ce chapitre est plus près des applications que les précédents, nous utiliserons la terminologie classique du calcul des probabilités et de l'analyse statistique.

§ 1. Position du problème

Soit X un espace métrique séparable, $X_1, X_2, \ldots, X_n, \ldots$ une suite d'éléments aléatoires à valeurs dans X admettant tous une même loi de probabilité : la mesure m de classe M_0 définie sur un σ-anneau de parties B de X.

Soit $\varphi(e, x)$ la caractéristique de l'ensemble $e \in B$:

$$\varphi(e, x) = 1 \qquad \text{si } x \in e,$$
$$\varphi(e, x) = 0 \qquad \text{si } x \notin e.$$

Pour chaque x, $\varphi(e, x)$ est une mesure définie sur B : c'est la mesure qui correspond à une masse unité placée au point x. Si au lieu d'un x certain, on considère un élément Y aléatoire à valeurs dans X, $\varphi(e, Y)$ est un élément aléatoire à valeurs dans l'espace des mesures sur X.

Posons :

$$m_n(e) = 1/n \sum_{i=1}^{n} \varphi(e, X_i);$$

$m_n(e)$ est encore un élément aléatoire dans l'espace des mesures. Si les X_i sont mutuellement indépendants, le problème de la convergence de la répartition empirique vers la répartition théorique est celui de la convergence stochastique des mesures aléatoires $m_n(e)$ vers la mesure certaine m.

R. Fortet et E. Mourier observent alors que le problème est ramené à une étude de lois des grands nombres. Il suffit, en effet, de trouver un espace de Banach séparable M contenant toutes les mesures, combinaisons linéaires de $\varphi(e, x)$,

$$\sum_{j=1}^{n} a_j \varphi(e, x_j),$$

pour que l'étude du comportement asymptotique de $m_n(e)$ se ramène à l'étude de lois des grands nombres pour des éléments aléatoires prenant leurs valeurs dans l'espace M. Mlle Mourier ayant obtenu dans sa thèse [26] des lois des grands nombres pour des éléments aléatoires satisfaisant à 3 conditions que nous préciserons plus loin et à valeurs dans un espace de Banach séparable, le problème est alors résolu. M. R. Fortet et Mlle E. Mourier ont effectivement trouvé un espace de Banach séparable répondant à la question et ont donc obtenu les résultats suivants :

Théorème 1

Presque sûrement $|m_n(e) - m(e)|$ tend vers zéro lorsque $n \to \infty$.

Théorème 2

$E(|m_n(e) - m(e)|)$ tend vers zéro lorsque $n \to \infty$. Dans ces

deux énoncés, les normes sont prises comme normes dans l'espace de Banach M choisi.

Les espaces M choisis par M. R. Fortet et Mlle E. Mourier sont les suivants :

1°) Si m est telle que :

$$m((x_0, x)) < +\infty$$

où x_0 est un point fixe quelconque de X et où (x_0, x) désigne la distance dans l'espace X, métrique par hypothèse, on adopte l'espace M défini comme suit :

Soit F l'espace des fonctions numériques $f(x)$ définies sur X et telles que :

$$\sup_{x', x'' \in X} \frac{|f(x') - f(x'')|}{(x', x'')} < +\infty.$$

C'est un espace de Banach avec la norme :

$$|f| = \sup_{x', x'' \in X} \frac{|f(x') - f(x'')|}{(x', x'')}.$$

Soit F* le dual fort de F.

D'autre part X étant séparable, soit (y_i) une suite dénombrable de points de X, dense dans X. Considérons les fonctions d'ensemble du type :

$$\varphi(e) = \varphi(e, y_j) - \varphi(e, y_k),$$

où y_j et y_k sont dans la suite (y_i).

Il est clair que $\varphi(e)$ est un élément de F ; en effet :

$$< \varphi(e), f > = f(x)\, d\, \varphi(e) = f(y_j) - f(y_k).$$

On prend alors pour espace M le plus petit sous-espace vectoriel fermé de F qui contienne tous les $\varphi(e)$.

2°) Dans le cas général, on considère l'espace G des fonctions de F telles que :

$$\sup_{x \in X} |f(x)| = N(f) < +\infty$$

et on prend sur G la norme

$$||f|| = |f| + N(f).$$

L'espace M sera un sous-espace séparable convenable du dual G* de G.

§ 2. Méthode générale

Supposons que les mesures $\varphi(e, x)$ et $m(e)$ soient considérées comme appartenant à un espace de Banach de mesures M, construit comme dual d'un espace vectoriel de fonctions continues U. Supposons en outre que l'on connaisse les fonctionnelles linéaires sur M, c'est-à-dire le bidual de U et que ce bidual soit formé d'une certaine classe de fonctions sommables pour toute mesure de M, c'est-à-dire que toute fonctionnelle linéaire sur M soit du type $m(f)$, où f est une fonction sommable (m), et où m varie dans M.

I. Observons d'abord que dans ce cadre général les résultats de M. R. Fortet et Mlle E. Mourier [13] sont encore valables. En effet, les éléments aléatoires :

$$m_n(e) = 1/n \sum_{r=1}^{n} [\varphi(e, X_i) - m(e)] = 1/n \sum_{i=1}^{n} m_0(e, X_i)$$

sont dans le sous-espace de M séparable engendré par $m(e)$ et les $\varphi(e, y_n)$, où les y_n sont une suite dénombrable partout dense dans X, (ce qui existe puisque, rappelons-le, X est supposé séparable). Et d'autre part les trois conditions du théorème de E. Mourier sont bien vérifiées :

1°) Pour tout m^* du dual de M,

$$< m^*, m_0(e, x) >$$

est une fonction mesurable de x. En effet, c'est une fonction continue, car,

$$\begin{aligned} &|<m^*, m_0(e, x')> - <m^*, m_0(e, x'')>| = \\ &= |<m^*, m_0(e, x') - m_0(e, x'')>| \\ &= |<m^*, \varphi(e, x') - \varphi(e, x'')>| \\ &\leqslant |m^*|| \varphi(e, x') - \varphi(e, x'')|. \end{aligned}$$

Or les mesures $\varphi(e, x') - \varphi(e, x'')$ ont une norme de l'ordre de (x', x''), puisqu'on obtient cette norme en appliquant les mesures aux fonctions de U qui sont continues et on a :

$$|\varphi(e, x') - \varphi(e, x'')| = \sup_{|f|\leqslant 1} |[\varphi(e, x') - \varphi(e, x'')]\,(f(x))|$$
$$= \sup_{|f|\leqslant 1} |f(x') - f(x'')|;$$

ce dernier terme est arbitrairement petit avec (x', x'') puisque f est continue.

Il en résulte que $< m^*, m_0(e, x) >$ est bien continue et, donc, mesurable.

2°) Montrons maintenant que la moyenne de $m_0(e, X_i)$ est finie. Observons que :

$$|m_0(e, X_i)| = \sup_{|u|\leqslant 1} |\varphi(e, X_i)\,(u(x)) - m(u(x))|$$
$$= \sup_{|u|\leqslant 1} |u(X_i) - m(u(x)).$$

Donc la moyenne cherchée est :

$$\mathrm{E}\,(m_0(e, X_i)) = \sup_{|u|\leqslant 1} \int |u(X_i) - m(u(x))|\, dm(e)$$
$$\leqslant \sup_{|u|\leqslant 1} \int |u(X_i)|\, |dm|(e) + \sup |m(u(x))|$$
$$\leqslant 2 \sup_{|u|\leqslant 1} m(u(x)),$$

qui est bien finie puisque c'est la norme de m comme élément de M.

3°) Montrons enfin que la moyenne de $m_0(e, X_i)$ est nulle.

Par définition de l'intégrale de Pettis, on a, pour toute fonctionnelle linéaire m^* sur l'espace M,

$$\mathrm{E}\,(< m^*, m_0(e, X_i) >) = < m^*, \mathrm{E}(m_0(e, X_i)) >.$$

Or toute fonctionnelle m^* est déterminée par une fonction $f(x)$, sommable pour toute mesure de M. On a donc :

$$< m^*, \mathrm{E}(m_0(e, X_i)) > = \mathrm{E}\,(\int f(x)\, dm_0(e, X_i))$$
$$= \mathrm{E}\,(f(X_i) - m(f))$$
$$= m(f) - m(f) = 0.$$

Puisque ceci est vrai pour tout m^* de M*, $\mathrm{E}(m_0(e, X_i)) = 0$.

Les trois conditions d'applicabilité de la loi des grands nombres sont donc vérifiées dans ce cas général et on a, comme dans R. Fortet et E. Mourier, que :

Théorème. Presque-sûrement $|m_n(e) - m(e)|$ tend vers zéro lorsque $n \to \infty$.

La norme est prise dans ce théorème au sens d'un espace M du type défini au début de ce paragraphe.

2. Mais, étant en possession du théorème central limite pour des espaces de Banach quelconques, nous pouvons compléter ce résultat.

Considérons l'élément aléatoire à valeur dans M :

$$\sqrt{n}\,[m_n(e) - m(e)] = 1/\sqrt{n} \sum_{i=1}^{n} m_0(e, X_i).$$

La loi de probabilité de cet élément tend, en vertu du théorème central limite, vers une loi laplacienne.

Nous allons exprimer cette loi laplacienne par sa caractéristique. Comme on sait que $E(m_0(e, X_i)) = 0$, cette caractéristique est de la forme

$$\varphi\,(m^*) = \exp\,[-1/2\,E\,(< m^*, M >^2)]$$

où m^* est un élément de M*, et M parcourt espace de mesure M. La moyenne E est prise par rapport à la mesure correspondant à l'élément aléatoire $m_0(e, X_i)$.

Comme tout élément de M est déterminé par une fonction sommable pour toute mesure $f(x)$, on a :

$$\varphi(m^*) = \exp\,[-1/2\,E\,\{M(f(x))^2\}].$$

Quelle est la loi de l'élément aléatoire $m_0(e, X_i)$?

On a :

$$m_0(e, X_i) = \varphi(e, X_i) - m(e).$$

Les seuls ensembles de M qui aient une probabilité nulle sont donc ceux qui contiennent une mesure δ_x ou $m(e)$ et on a :

$$\text{prob}\,[\varphi\,(e, X_i)\,E\,\{\,\delta_x : x \in e\}] = m(e),$$

puisque X_i a pour loi $m(e)$.

Il vient donc :

$$\varphi(m^*) = \exp\,[-1/2\,E\,\{\int M(f(x))^2\,d\,\delta_i\} - \int M(f(x))^2\,d\,\delta_{m(e)}]$$

où δ_i et $\delta_{m(e)}$ désignent des mesures sur l'espace M, formées de la masse unité, placée respectivement à l'élément δ_{x_i} et à l'élément $m(e)$ de M; le signe E de moyenne subsistant dans le premier terme de l'exposant désigne cette fois une moyenne par rapport à l'élément aléatoire X_i.

Les intégrales étant prises pour des masses unités, on les effectue en remplaçant simplement la variable M par l'élément voulu; il vient :

$$\varphi(m^*) = \exp\,[-1/2\,\mathrm{E}\,\{\,\delta_{x_i}\,f(x))^2\} - m(f(x))^2]$$
$$= \exp\,[-1/2\,\mathrm{E}\,\{\,f(\mathrm{X}_i)^2\} - m(f(x))^2].$$

Ou, puisque la loi de X_i est $m(e)$:

$$\varphi(m^*) = \exp\,[-1/2\,m(f(x)^2) - m(f(x))^2].$$

Nous avons donc obtenu, par sa caractéristique, la loi limite de

$$\sqrt{n}\,[m_n(e) - m(e)],$$

considéré comme élément aléatoire à valeurs dans M, pour $n \to \infty$.

Observons cependant que $m(f(x)^2)$ et $m(f(x))$ doivent être estimés dans cette expression à partir des données. Pour pouvoir aller plus loin nous allons dans le paragraphe suivant particulariser l'espace M.

§ 3. Résultat complémentaire.

Nous allons maintenant décrire un procédé qui permet de plonger les mesures $m_n(e)$ et $m(e)$ dans un espace de Hilbert.

Soit U_0 l'espace des fonctions continues et bornées. Nous allons munir U_0 d'une topologie qui en fait un espace préhilbertien, (c'est-à-dire un espace dont la fermeture est un espace de Hilbert). Pour cela, il suffit de définir un produit scalaire convenable sur U_0.

X étant séparable, soit (x_i) une suite de points de X, dénombrable et partout dense dans X. Soit (a_i) une suite de nombres strictement positifs de somme convergente. Supposons par exemple que :

$$\sum_{i=1}^{\infty} a_i = 1.$$

Le produit scalaire que nous introduisons sur U_0 est alors le suivant :

$$(f, g) = \int f(x)\,g(x)\,dm(e) + \sum_{i=1}^{\infty} a_i\,f(x_i)\,g(x_i) \qquad (1)$$

ce produit est bien défini puisque f et g, étant dans U_0, sont bornées, donc de carré sommable pour la mesure $m(e)$; la série prise comme second terme est bien convergente avec Σa_i puisque f et g sont bornées.

La linéarité et les autres propriétés du produit scalaire hilbertien sont immédiates d'après la définition.

Soit maintenant $\overline{U_0}$ l'espace de Hilbert, fermeture de l'espace U_0 pour la topologie définie par le produit scalaire (1). Soit $\overline{U_0}^*$ le dual de cet espace de Hilbert. Il est clair qu'une fonctionnelle linéaire continue sur $\overline{U_0}$ est complètement déterminée par ses valeurs sur U_0 puisque, par définition de la fermeture, U_0 est dense dans $\overline{U_0}$.

Je dis que $m(e)$ et chacun des $\varphi(e, x)$ pour tout x de X sont des fonctionnelles linéaires continues sur U_0.

En effet, la linéarité est évidente. Il suffit de montrer la continuité.

a) Si une suite de fonctions converge dans U_0 préhilbertien, elle converge en moyenne d'ordre 2 par rapport à la mesure $m(e)$ à cause du premier terme du produit scalaire. Elle converge donc a fortiori en moyenne (m), c'est-à-dire que $m(e)$ est bien continue.

b) Soit $\varphi(e, x_n)$ avec x_n dans la suite partout dense choisie dans X. Il est clair que la fonctionnelle linéaire $\varphi(e, x_n)$ est continue. Cela résulte de la présence du terme $a_n(fx_n)g(x_n)$ dans le produit scalaire (1). En effet, ce terme entraîne que si une suite de fonctions converge dans U_0 préhilbertien, elle converge au point x_n et donc converge au sens de la mesure $\varphi(e, x_n)$ qui est une masse unité placée en x_n.

c) Mais si la suite de fonctions converge en tout point de l'ensemble (x_i), elle converge en tout point de X. Car (x_i) est partout dense dans X et les fonctions de U_0 sont continues. Donc $\varphi(e, x)$ est bien une fonctionnelle linéaire continue pour tout x de X.

On peut donc appliquer les résultats du paragraphe précédent. L'expression

$$\sqrt{n}\,[m_n(e) - m(e)] \tag{2}$$

est un élément aléatoire à valeur dans l'espace de Hilbert $\overline{U_0}^*$ et sa loi, donnée par une mesure sur $\overline{U_0}^*$, a pour fonction caractéristique :

$$\varphi(m^*) = \exp\,[-1/2\, m(f(x)^2) - m(f(x))^2] \tag{3}$$

Un procédé employé en analyse statistique pour préciser la connaissance de l'élément (2), consiste à chercher la loi de la variable aléatoire, carré de la norme de l'élément aléatoire ayant pour caractéristique la fonction (3). Mais puisque notre élément aléatoire (2) prend maintenant ses valeurs dans un espace de Hilbert, on connaît cette loi. Elle est donnée par un théorème de R. Fortet et E. Mourier [12].

Théorème :

Si Y est un élément aléatoire laplacien prenant ses valeurs dans un espace de Hilbert séparable et de moyenne nulle, la fonction caractéristique de la variable aléatoire $|Y|^2$ est égale à :

$$\varphi(v) = [D(2iv)]^{-1/2}$$

où $D(\lambda)$ désigne le déterminant de Fredholm de la forme quadratique complètement continue

$$E(<x^*, U>^2),$$

de x^* dans le dual de l'espace de Hilbert.

Or ici $\sqrt{n}\,[m_n(e) - m(e)]$ prend ses valeurs dans le sous-espace de Hilbert fermé engendré par les $\varphi(e, x_i)$ et $m(e)$, sous-espace qui est donc séparable. Et de plus on a vu que (2) a bien une moyenne nulle, (cfr. § 2).

Nous avons donc dans le cas général d'éléments aléatoires à valeurs dans un espace métrique séparable X non seulement la caractéristique de l'élément aléatoire

$$\sqrt{n}\,[m_n(e) - m(e)]$$

mais aussi la caractéristique du carré de sa norme. Ces résultats n'étaient connus que dans le cas où X est la droite réelle.

BIBLIOGRAPHIE

[1] ALEXANDROFF P., *Topologie*, Berlin, Springer, (1935).

[2] BANACH S., *Théorie des opérations linéaires*, Varsovie, (1932).

[3] BOCHNER, Integration von Functionen deren Werte die Elemente eines Vectorraumes sind, *Fund. Math.*, 20, (1933).

[4] BOURBAKI N., *Eléments de Mathématique*, Livre III, *Topologie*, Livre V, *Espaces vectoriels topologiques*, Livre VI, *Intégration*, Paris, Hermann, (1940-).

[5] BOURBAKI N., Sur certains espaces vectoriels topologiques, *Annales de l'Institut Fourier*, Grenoble, 2, (1950), 5-16.

[6] CHOQUET, Convergences, *Annales de l'Université de Grenoble*, 23, (1948), p. 25.

[7] DARMOIS G., *Comptes rendus, Ac. Sc. Paris*, 232, (1951), 1999-2000.

[8] DIEUDONNÉ J., Sur le théorème de Radon-Nikodym, *Annales de l'Université de Grenoble*, 23, (1948).

[9] DIEUDONNÉ J. et SCHWARTZ L., La dualité dans les espaces (F) et (LF). *Annales de l'Institut Fourier*, Grenoble, I, (1949), 61-102.

[10] DOSS, S., Sur la moyenne d'un élément aléatoire dans un espace distancié, *Bull. Sciences Math.*, 73, (1949).

[11] DOSS S., Sur la convergence stochastique dans les espaces uniformes, *Annales de l'école normale supérieure*, (1954), 87-100.

[12] FORTET R. et MOURIER E., Résultats complémentaires sur les éléments aléatoires prenant leurs valeurs dans un espace de Banach, *Bull. Sc. Math.*, 78, (1954), 1-16.

[13] FORTET R. et MOURIER E., Convergence de la répartitition empirique vers la répartition théorique, *Annales de l'Ecole normale supérieure*, (1954), 267-285.

[14] FRÉCHET M., Les éléments aléatoires de nature quelconque dans un espace distancié, *Annales Inst. H. Poincaré*, 10, (1948), 215-310.

[15] FRÉCHET M., Généralisation de la loi de probabilité de Laplace, *Annales Inst. H. Poincaré*, 12, (1951), 1-29.

[16] HALMOS P. R., *Measure Theory*, Van Nostrand, New-York, (1950).

[17] HALMOS P. R., AMBROSE W. et KAKUTANI S., The decomposition of measure, *Duke J. Math.*, 9, (1942), 43-47.

[18] HAUSDORFF F., *Grundzüge der Mengenlehre*, Leipzig, (1914).

[19] HILLE E., Functional Analysis and Semi-groups, *Amer. Math. Soc.*, (1948).

[20] KAKUTANI S., Concrete representation of abstract L space and the mean ergodic theorem, *Annals of Math.*, 42, (1941), 523-537.

[21] KURATOWSKI, *Topologie*, Tome I et II, (3r édition), 1952.

[22] LEBESGUE H., *Leçons sur la théorie de l'intégration*, Gauthier-Villars, Paris, (1928).

[23] LE CAM L., Un instrument d'étude des fonctions aléatoires : la fonctionnelle caractéristique. *Comptes rendus Ac. Sc. Paris*, 224, (1947), 710-711.

[24] MACKEY G. W., On infinite dimensional linear spaces, *Trans. Amer. Math. Soc.*, 57, (1945), 155-207.

[25] MACKEY G W., On convex topological linear spaces, *Trans. Amer. Math. Soc.*, 60, (1946), 520-537.

[26] MOURIER E., *Eléments aléatoires dans un espace de Banach*, Thèse. Paris. (1952).

[27] PETTIS, On integration in vector spaces, *Trans. Amer. Math. Soc.*, 44, (1948).

[28] PROHOROV, Distributions de probabilités dans des espaces fonctionnels. *Uspehi Math. Nauk.* (N. S.) 8, (1953), 165-167.

[29] RIESZ F., Sur les opérations fonctionnelles linéaires, *Comptes rendus Ac. Sc.*, 149, (1909), 974-976.

[30] STONE M. H., The theory of representations of Boolean algebras, *Trans. Amer. Math. Soc.*, 40, (1936), 37-111.

[31] STONE M. H., Application of the theory of Boolean rings to general topology, *Trans. Amer. Math. Soc.*, 41, (1937), 375-481.

[32] DE LA VALLÉE-POUSSIN C., *Intégrale de Lebesgue. Fonctions d'ensemble. Classes de Baire.* Gauthier-Villars, Paris, (1934).

[33] YOSIDA et KAKUTANI, Operator theorical treatment of Markoff process and mean ergodic theorem. *Annals Math.*, 42, (1941).

[34] WEIL A., L'intégration dans les groupes topologiques et ses applications. Paris, Hermann, (2r édit.), (1953).

Deuxième Section

Sciences Physiques et Chimiques

M. Ch. Manneback fait une communication orale, sur les molécules de benzène et de naphtalène.

(26.977) Etablissements Ceuterick, s.c., 66 rue V. Decoster, Louvain.
Resp. L. Pitsi, 25, rue Dagobert, Louvain (Belgique).

Revue des Questions Scientifiques

Cette revue, fondée en 1877 par la Société scientifique de Bruxelles, se compose actuellement de cinq séries : la **première série** comprend 30 volumes (1877-1891); la **deuxième**, 20 volumes (1892-1901); la **troisième**, 30 volumes (1902-1921); la **quatrième**, 30 vol. (1922-1936). La livraison de janvier 1937 inaugura la **cinquième série**.

La revue fut interrompue par la guerre, après le deuxième fascicule, avril 1940, du tome 117 de la collection. Ce tome fut achevé par les fascicules de février et juillet 1946.

Depuis 1947, paraît chaque année un volume, en quatre fascicules d'environ 160 pages chacun, 20 janvier, 20 avril, 20 juillet, 20 octobre.

Depuis 1947 la Revue est aussi l'organe de l'UNION CATHOLIQUE DES SCIENTIFIQUES FRANÇAIS.

Administration et Rédaction : **11, rue des Récollets, Louvain** (Belgique)

PRIX D'ABONNEMENT

Belgique, Congo belge et Grand-Duché de Luxembourg	200 francs belges
abonnement de soutien	500 francs belges
abonnement d'honneur	1000 francs belges
France et Union française (s'adresser à la librairie Masson, 120 boulevard St Germain, Paris 6e)	1400 francs franç.
Autres pays (s'adresser au Secrétariat, à Louvain)	280 francs belges

PUBLICATIONS DE LA SOCIÉTÉ SCIENTIFIQUE

EXTRAIT DU CATALOGUE

ANNALES DE LA SOCIETE SCIENTIFIQUE DE BRUXELLES,

t. I à XLVI, 1875 à 1926. Chaque vol. in-8° de 400 à 600 pages . . F 150,00

ANNALES DE LA SOCIETE SCIENTIFIQUE DE BRUXELLES,

Série A (sc. mathématiques), t. XLVII à LVI (1927 à 1936) . . . . . F 70,00
Série B (sc. physiques et naturelles) » » » F 70,00
Série C (sc. médicales) (1927 et 1928) F 100,00 — (1929 à 1933) (1931 à 1936) F 40,00
Série D (sc. économ. et techniques) (1927 à 1929) F 20,00 — (1930) (1931 à 1936) F 60,00
Série I (sc. mathématiques et physiques), tt. LVII à LXVII (1937 à 1953) F 100,00
tt. LXVIII et LXIX (1954 et 1955) F 100,00
Série II (sc. naturelles et médicales), tt. LVII à LX (1937 à 1940, 46) F 70,00
Série III (sc. économiques), tt. LVII à LX (1937 à 1940, 46) . . . . . F 100,00

REVUE DES QUESTIONS SCIENTIFIQUES,

t. I à XCII (1877 à 1927). Les deux volumes annuels . . . . . F 200,00
Le fascicule trimestriel . . . . . F 60,00
t. XCIII à CXVI (1928 à 1939). Les deux volumes annuels . . . . F 200,00
Le fascicule . . . . . F 40,00
t. CXVII (1940 et 1946), tt. CXVIII à CXXVI (1947 à 1955). Le volume F 200,00
Le fascicule . . . . . F 60,00

TABLES ANALYTIQUES DES ANNALES,

t. I à XXV (1875 à 1901) . . . . . F 20,00
t. XXVI à XLVI (1902 à 1926) . . . . . F 40,00

TABLES ANALYTIQUES DE LA REVUE,

t. I à L (1877 à 1901). . . . . F 20,00
t. LI à LXXX (1902 à 1921) . . . . . F 20,00
t. LXXXI à CX (1922 à 1936) . . . . . F 30,00

MONOGRAPHIES DE SCIENCES NATURELLES

I — **B. Tougarinoff.** Les réactions organiques dans l'analyse qualitative minérale (cations). — Un vol. in-8° de 107 pages (1930) : en Belgique, F 24,00; autres pays : F 30,00.

II — **V. Schaffers.** Le paratonnerre et ses progrès récents. Un vol. in-8° de 90 pages (1931) : en Belgique, F 24,00; autres pays : F 30,00.

IV — **F. Kaisin et E. de Pierpont.** — Hydrogéologie des Calcaires de la Belgique. Un vol. in-8° de 111 pages, avec 35 fig. et un plan hors texte (1939) : en Belgique, F 24,00; autres pays, F 30,00.

MONOGRAPHIES MEDICALES

I — **M. Schillings.** Le rein en fer à cheval. Un vol. in-8° de 104 pages, avec 8 planches hors-texte (1938) : en Belgique, F 70,00; autres pays, F 90.00.

III — **P. Van Gehuchten.** La pathologie du système pallido-strié. Un vol. in-8° de 52 pages, avec 8 planches hors-texte (1930) : en Belgique, F 24,00; autres pays, F 30,00.

MONOGRAPHIES DES SCIENCES ECONOMIQUES

I — **A. Henry.** La structure technique de l'agriculture belge et ses particularités en Wallonie et en Flandre. Un vol. de 66 pages . . . . . F 20,00

II — **A. Henry.** Les variations régionales de l'Agriculture en Belgique. Un vol. de 50 pages . . . . . F 10,00

III — **A. Delpérée.** La réglementation conventionnelle des conditions de travail en Belgique. Un vol. de 200 pages . . . . . F 60,00